

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 05:35:31 ; Search time 172 Seconds
(without alignments)
1072.585 Million cell updates/sec

Title: "": US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPEETQPLGPPGSA.....CVPETKGTLEQITAHFEGR 477

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:*
1: Geneseqp1980s:*
2: Geneseqp1990s:*
3: Geneseqp2000s:*
4: Geneseqp2001s:*
5: Geneseqp2002s:*
6: Geneseqp2003as:*
7: Geneseqp2003bs:*
8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2457	100.0	477	AAE04888	Aae04888 Human tra
2	2457	100.0	477	ABP58364	Abp58364 Human sol
3	2448	99.6	477	AAB66932	Aab66932 Human GLU
4	2224.5	90.5	478	AAB66939	Aab66939 GLUTX1 co
5	2162.5	88.0	478	AAB66933	Aab66933 Rat GLUTX
6	2150	87.5	477	AAB66934	Aab66934 Murine GL
7	2067.5	84.1	408	ABM84862	Abm84862 Human dia
8	1912.5	77.8	382	ABM84863	Abm84863 Human dia
9	1664	67.7	353	ADL33342	Adl33342 Human tra
10	1491	60.7	326	ABB89717	Abb89717 Human pro
11	1135	46.2	262	AAE06579	Aae06579 Human pro
12	1131	46.0	248	ABU11283	Abu11283 cDNA enco
13	969.5	39.5	503	AAB66938	Aab66938 Rat GLUTX
14	953	38.8	507	AAB66941	Aab66941 GLUTX3 co
15	948	38.6	507	AAB66937	Aab66937 Human GLU
16	948	38.6	507	AAW93417	Aaw93417 Human pol
17	948	38.6	507	ADA84077	Ada84077 Human SLC
18	948	38.6	507	ABP58365	Abp58365 Human sol
19	948	38.6	507	ABP58363	Abp58363 Human sol
20	948	38.6	507	ADL31001	Adl31001 Human pro
21	948	38.6	507	ADR14109	Adr14109 Human NF-
22	894	36.4	445	AAM93364	Aam93364 Human pol
23	894	36.4	445	ADJ30894	Adj30894 Human pol
24	846	34.4	431	ADP29770	Adp29770 Human sec
25	808.5	32.9	470	ABM83266	Abm83266 Human dia

26	788.5	32.1	489	4	ABB63511	Abb63511 Drosophil
27	788	32.1	539	4	ABB65350	Abb65350 Drosophil
28	764.5	31.1	433	4	ABB63758	Abb63758 Drosophil
29	727	29.6	465	4	ABB62408	Abb62408 Drosophil
30	723	29.4	501	8	ADG47942	Adg47942 Beta-vulg
31	705	28.7	497	4	ABB57922	Abb57922 Drosophil
32	670	27.3	447	3	AAG13749	Aag13749 Arabidops
33	670	27.3	463	3	AAG13748	Aag13748 Arabidops
34	669.5	27.2	437	3	AAG13750	Aag13750 Arabidops
35	666.5	27.1	444	4	ABB58057	Abb58057 Drosophil
36	662	26.9	447	3	AAG20977	Aag20977 Arabidops
37	662	26.9	463	3	AAG20976	Aag20976 Arabidops
38	661.5	26.9	437	3	AAG20978	Aag20978 Arabidops
39	658.5	26.8	553	8	ADG47940	Adg47940 Corn Beta
40	652.5	26.6	448	3	AAG38870	Aag38870 Arabidops
41	652.5	26.6	464	3	AAG38869	Aag38869 Arabidops
42	652	26.5	438	3	AAG38871	Aag38871 Arabidops
43	647	26.3	491	4	ABB65373	Abb65373 Drosophil
44	647	26.3	491	4	ABB57923	Abb57923 Drosophil
45	639	26.0	481	8	ADM48307	Adm48307 Polypepti

ALIGNMENTS

RESULT 1
AAE04888
ID AAE04888 standard; protein; 477 AA.
XX AC AAE04888;
XX DT 10-SEP-2001 (first entry)
XX DE Human transporter and ion channel-1 (TRICH-1) protein.
XX KW Human; transporter and ion channel-1; TRICH-1; vaccine; cystic fibrosis;
KW gene therapy; amyotrophic lateral sclerosis; amnesia; muscular dystrophy;
KW hypertension; angina; neurological disorder; asthma; bipolar disorder;
KW dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia;
KW Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke;
KW Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;
KW demyelinating disease; mental disorder; schizophrenia; polymyositis;
KW muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;
KW dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;
KW rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;
KW sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;
KW scleroderma; pulmonary artery stenosis; neutropenic; Addison's disease;
KW malabsorption syndrome; hypercholesterolaemia; cancer.

Key	Domain	Location/Qualifiers
FT	Domain	29..474
FT	Domain	/note= "Sugar transporter domain"
FT	Domain	259..279
FT	Domain	/label= Transmembrane_domain
FT	Domain	293..313
FT	Domain	/label= Transmembrane_domain
FT	Domain	320..339
FT	Domain	/label= Transmembrane_domain
FT	Domain	438..457
FT	Domain	/label= Transmembrane_domain

WO200146258-A2.

28-JUN-2001.

22-DEC-2000; 2000WO-US035095.

23-DEC-1999; 99US-0172000P.

14-JAN-2000; 2000US-0176083P.

21-JAN-2000; 2000US-0177332P.

28-JAN-2000; 2000US-0178572P.

PR 02-FEB-2000; 2000US-0179758P.
 PR 10-FEB-2000; 2000US-0181625P.
 XX (INCY-) INCYTE GENOMICS INC.
 XX Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R;
 PI Lal P, Hillman JL, Azimzai Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;
 PI Tang YT, Khan FA;
 XX WPI; 2001-418042/44.
 DR N-PSDB; AAD09552.
 XX Novel human transporter and ion channel proteins useful for treating and
 PT preventing transport, neurological, muscle and immunological disorders.
 XX Claim 1; Page 112-113; 160pp; English.
 XX The present sequence is transporter and ion channel-1 (TRICH-1) protein.
 CC TRICH is used as vaccine. TRICH is useful for treating a disease or
 CC condition associated with decreased expression of functional TRICH, such
 CC as transport disorder including amyotrophic lateral sclerosis, cystic
 CC fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth disease,
 CC Duchenne muscular dystrophy, angina and hypertension, neurological
 CC disorders including Alzheimer's disease, amnesia, bipolar disorder,
 CC dementia, depression, epilepsy, ischaemic cerebrovascular disease,
 CC stroke, cerebral neoplasms, Pick's disease, Huntington's disease and
 CC Parkinson's disease, demyelinating diseases, mental disorders including
 CC mood, anxiety, schizophrenia and seasonal affective disorder, muscle
 CC disorder including cardiomyopathy, myocarditis, polymyositis,
 CC dermatomyositis, arrhythmias and asthma and immunological disorders
 CC including AIDS, adult respiratory distress syndrome (ARDS), allergies,
 CC anæmia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
 CC syndrome, systemic lupus erythematosus and other diseases including
 CC sickle cell anaemia, Wilson's disease, cataracts, infertility, pulmonary
 CC artery stenosis, Grave's disease, Cushing's disease, Addison's disease,
 CC glucose-galactose malabsorption syndrome, hypercholesterolaemia, cancers
 CC psoriasis and viral, bacterial, fungal, helminthic and protozoal
 CC infections. TRICH DNA is useful in gene therapy and in diagnostic
 CC purposes
 XX
 SQ Sequence 477, AA;
 Query Match 100.0%; Score 2457; DB 4; Length 477;
 Best Local Similarity 100.0%; Pred. No. 8.8e-227;
 Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MTPEDPEETQPLGPGGAPGRGRVFLAFAAALGFLSGFALGYSSPAIPSLQRAAPP 60
 DB 1 MTPEDPEETQPLGPGGAPGRGRVFLAFAAALGFLSGFALGYSSPAIPSLQRAAPP 60
 QY 61 APRLDDAASWFGAVVTLGAAAGVGLGWLVDVGRGKLSLLCSVPFVAGFAVITAAQDV 120
 DB 61 APRLDDAASWFGAVVTLGAAAGVGLGWLVDVGRGKLSLLCSVPFVAGFAVITAAQDV 120
 QY 121 WMLLGGRLLTGLACGVASLVAIPYIIEIAYPAVRGLLSCVQLMVVVGILLAYLAGWVLE 180
 DB 121 WMLLGGRLLTGLACGVASLVAIPYIIEIAYPAVRGLLSCVQLMVVVGILLAYLAGWVLE 180
 QY 181 WRWLAVLGCVPVPSIMLLLMCMFETPRFLITQHRROEAMAAALRFLWGSQGWEDDPGAE 240
 DB 181 WRWLAVLGCVPVPSIMLLLMCMFETPRFLITQHRROEAMAAALRFLWGSQGWEDDPGAE 240
 QY 241 QSFHLALLRPGYIKPFIIGVSLMAFQQLSGVNVAMFYAETIFEEAKFKDSSLASVVGV 300
 DB 241 QSFHLALLRPGYIKPFIIGVSLMAFQQLSGVNVAMFYAETIFEEAKFKDSSLASVVGV 300
 QY 301 IQVLFATAVALIMDRAGRLLLVLSGVWVVFSTSAFGAYFKLTQGGPGNSSHVAISAPVS 360
 DB 301 IQVLFATAVALIMDRAGRLLLVLSGVWVVFSTSAFGAYFKLTQGGPGNSSHVAISAPVS 360
 QY 361 AQPVDASVGLAWAVGSMCLFIAGFAVGWGPFWLLMSEIFPLHVKVATGICVLTNWL 420
 DB 361 AQPVDASVGLAWAVGSMCLFIAGFAVGWGPFWLLMSEIFPLHVKVATGICVLTNWL 420

QY 421 AFLVTKFSSIMVLRPVGAFWLASAFICISVLFLLFCVPETKGTLEQITAHFEGR 477
 DB 421 AFLVTKFSSIMVLRPVGAFWLASAFICISVLFLLFCVPETKGTLEQITAHFEGR 477
 RESULT 2
 ID ABP58364 standard; protein; 477 AA.
 XX AC ABP58364;
 XX 07-APR-2003 (first entry)
 DT Human solute carrier type 2A polypeptide 12735153 and 7657681.
 DE Human; solute carrier type 2A; SLC2A; glucose transporter; p53; cancer;
 KW cytosolic; gene therapy.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FT Domain 27..49 /note= "transmembrane domain"
 FT Domain 29..474 /note= "sugar transporter domain"
 FT Domain 69..91 /note= "transmembrane domain"
 FT Domain 98..117 /note= "transmembrane domain"
 FT Domain 127..149 /note= "transmembrane domain"
 FT Domain 156..178 /note= "transmembrane domain"
 FT Domain 183..205 /note= "transmembrane domain"
 FT Domain 257..279 /note= "transmembrane domain"
 FT Domain 294..313 /note= "transmembrane domain"
 FT Domain 320..342 /note= "transmembrane domain"
 FT Domain 368..390 /note= "transmembrane domain"
 FT Domain 403..425 /note= "transmembrane domain"
 FT Domain 440..462 /note= "transmembrane domain"
 XX WO200298467-A1.
 PD 12-DEC-2002.
 XX 03-JUN-2002; 2002WO-US017419.
 XX 05-JUN-2001; 2001US-0296076P.
 PR 10-OCT-2001; 2001US-0328605P.
 PR 15-FEB-2002; 2002US-0357253P.
 XX (EXEL-) EXELIXIS INC.
 XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
 WPI; 2003-201283/19.
 DR N-PSDB; ABZ24792, ABZ24793, ABZ24794.
 XX Identifying candidate p53 pathway modulating agent as therapeutic target
 PT for disorders of defective p53 function e.g. cancer, by assaying purified
 PT solute carrier family 2 (SLC2A) polypeptide or nucleic acid with a test
 agent.
 XX Claim 13; Page 50-52; 58pp; English.

CC The present sequence is that of human solute carrier type 2A (SLC2A)
 CC polypeptides 1273153 and 7657681. In the present invention, genetic
 CC screens were designed to identify modifiers of the p53 pathway in
 CC Drosophila in which p53 was overexpressed. Human orthologues
 CC (polynucleotides and polypeptides) of one such modifier were then
 CC identified, including the present polypeptides. SLC2As are glucose
 CC transporter proteins with sugar transporter domains. They are attractive
 CC drug targets for treatment of pathologies associated with a defective p53
 CC signalling pathway, such as cancer. The invention provides in vitro and
 CC in vivo methods of assessing SLC2A function. Modulation of an SLC2A or
 CC its binding partners is useful for understanding the association of the
 CC p53 pathway and its members in normal and disease conditions and for
 CC developing diagnostic and therapeutic modalities of p53-related
 CC pathologies. SLC2A-modulating agents that act by inhibiting or enhancing
 CC SLC2A expression, directly or indirectly, e.g. by affecting an SLC2A
 CC function such as transport or binding activity, can be identified using
 CC methods provided. Modulators include small molecules, nucleic acids,
 CC antibodies, antisense oligonucleotides and phosphothioate morpholino
 CC oligomers (claimed)
 XX
 SQ Sequence 477 AA;

Query Match 100.0%; Score 2457; DB 6; Length 477;
 Best Local Similarity 100.0%; Pred. No. 8.8e-227; Indels 0; Gaps 0;
 Matches 477; Conservative 0; Mismatches 0;
 QY 1 MTPDEPEETQPLGGPGSAPRRVFLAAFAAALGPLSFGFALGYSSPAISLQRAAPP 60
 DB 1 MTPDEPEETQPLGGPGSAPRRVFLAAFAAALGPLSFGFALGYSSPAISLQRAAPP 60
 QY 61 APRLDDAAASWFGAVVTGAAAGGVLGGWLVDRAGKLSLLCSVPFVAGFAVITAAQDV 120
 DB 61 APRLDDAAASWFGAVVTGAAAGGVLGGWLVDRAGKLSLLCSVPFVAGFAVITAAQDV 120
 QY 121 WMLGGRLLTGLACGVASLAVPVYISETIAYPAVRGLLGSVQLMVVVGILLAYLAGWILE 180
 DB 121 WMLGGRLLTGLACGVASLAVPVYISETIAYPAVRGLLGSVQLMVVVGILLAYLAGWILE 180
 QY 181 WRWLAVLGCVPSPSLMLLMCFMPETPRFLLTQHRROEAMAAALRFLMGSEQWEDPPIGAE 240
 DB 181 WRWLAVLGCVPSPSLMLLMCFMPETPRFLLTQHRROEAMAAALRFLMGSEQWEDPPIGAE 240
 QY 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEITFEAKFKDSSLASVVGV 300
 DB 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEITFEAKFKDSSLASVVGV 300
 QY 301 IQVLFATAALIMDRAGRELLLVLSGVVMVFSTSAFGAYFKLTQGGPGNSSHVAISAPVS 360
 DB 301 IQVLFATAALIMDRAGRELLLVLSGVVMVFSTSAFGAYFKLTQGGPGNSSHVAISAPVS 360
 QY 361 AQPVDASVGLAWLAVGSMCLFIAGFVAGVGPIPLWLMSEIFPLHVKGVATGICVLTNWLM 420
 DB 361 AQPVDASVGLAWLAVGSMCLFIAGFVAGVGPIPLWLMSEIFPLHVKGVATGICVLTNWLM 420
 QY 421 AFLVTKFSSLMELVRPYGAFWLASAFICFISVFLTFPCVPETKGTLEQITAHFEGR 477
 DB 421 AFLVTKFSSLMELVRPYGAFWLASAFICFISVFLTFPCVPETKGTLEQITAHFEGR 477

RESULT 3
 AAB66932
 ID AAB66932 standard; protein; 477 AA.
 XX
 AC AAB66932;
 XX
 DT 17-APR-2001 (first entry)
 XX
 DE Human GLUTX1.
 XX
 KW Human; GLUTX; gene therapy; vaccine; hexose transport modulator;
 KW hexose transport disorder; ischaemia; diabetes; hypoglycaemia;
 KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
 XX

OS Homo sapiens.
 XX WO200104145-A2.
 XX
 PD 18-JAN-2001.
 XX
 PF 14-JUL-2000; 2000WO-IB001042.
 XX
 PR 14-JUL-1999; 99US-0143907P.
 PR 27-AUG-1999; 99US-0151140P.
 PR 23-FEB-2000; 2000US-0184285P.
 PR 13-JUL-2000; 2000US-00616132.
 XX
 (UYLA-) UNIV LAUSANNE.
 PA
 PI Thorens B, Ibberson M, Uldry M;
 XX
 XX WPI; 2001-112615/12.
 DR N-P8DB; AAF55865.
 XX
 PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
 PT ischaemia and diabetes.
 XX
 PS Claim 11; Page 70-71; 124pp; English.
 XX
 CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hypoglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is human GLUTX1
 XX
 SQ Sequence 477 AA;
 Query Match 99.6%; Score 2448; DB 4; Length 477;
 Best Local Similarity 99.6%; Pred. No. 6.4e-226;
 Matches 475; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
 QY 1 MTPDEPEETQPLGGPGSAPRRVFLAAFAAALGPLSFGFALGYSSPAISLQRAAPP 60
 DB 1 MTPDEPEETQPLGGPGSAPRRVFLAAFAAALGPLSFGFALGYSSPAISLQRAAPP 60
 QY 61 APRLDDAAASWFGAVVTGAAAGGVLGGWLVDRAGKLSLLCSVPFVAGFAVITAAQDV 120
 DB 61 APRLDDAAASWFGAVVTGAAAGGVLGGWLVDRAGKLSLLCSVPFVAGFAVITAAQDV 120
 QY 121 WMLGGRLLTGLACGVASLAVPVYISETIAYPAVRGLLGSVQLMVVVGILLAYLAGWILE 180
 DB 121 WMLGGRLLTGLACGVASLAVPVYISETIAYPAVRGLLGSVQLMVVVGILLAYLAGWILE 180
 QY 181 WRWLAVLGCVPSPSLMLLMCFMPETPRFLLTQHRROEAMAAALRFLMGSEQWEDPPIGAE 240
 DB 181 WRWLAVLGCVPSPSLMLLMCFMPETPRFLLTQHRROEAMAAALRFLMGSEQWEDPPIGAE 240
 QY 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEITFEAKFKDSSLASVVGV 300
 DB 241 QSFHLALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEITFEAKFKDSSLASVVGV 300
 QY 301 IQVLFATAALIMDRAGRELLLVLSGVVMVFSTSAFGAYFKLTQGGPGNSSHVAISAPVS 360
 DB 301 IQVLFATAALIMDRAGRELLLVLSGVVMVFSTSAFGAYFKLTQGGPGNSSHVAISAPVS 360
 QY 361 AQPVDASVGLAWLAVGSMCLFIAGFVAGVGPIPLWLMSEIFPLHVKGVATGICVLTNWLM 420
 DB 361 AQPVDASVGLAWLAVGSMCLFIAGFVAGVGPIPLWLMSEIFPLHVKGVATGICVLTNWLM 420
 QY 421 AFLVTKFSSLMELVRPYGAFWLASAFICFISVFLTFPCVPETKGTLEQITAHFEGR 477
 DB 421 AFLVTKFSSLMELVRPYGAFWLASAFICFISVFLTFPCVPETKGTLEQITAHFEGR 477

```

RESULT 4
ID AAB66939 standard; protein; 478 AA.
AC AAB66939;
DT 17-APR-2001 (first entry)
XX GLUTX1 consensus sequence.
DE GLUTX; gene therapy; vaccine; hexose transport modulator; human; rat;
KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia; murine;
KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
OS Homo sapiens.
OS Mus sp.
OS Rattus sp.
XX
XX WO200104145-A2.
XX
XX 18-JAN-2001.
XX
XX 14-JUL-2000; 2000WO-IB001042.
XX
XX 14-JUL-1999; 99US-0143907P.
XX
XX 27-AUG-1999; 99US-0151140P.
XX
XX 23-FEB-2000; 2000US-0184285P.
XX
XX 13-JUL-2000; 2000US-00616132.
XX
XX (UYLA-) UNIV LAUSANNE.
XX
XX Thorens B, Ibberson M, Uldry M;
XX
XX WPI; 2001-112615/12.
XX
XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
XX prevention, diagnosis and treatment of hexose transport disorders, e.g.
XX ischemia and diabetes.
XX
XX Claim 11; Page 74-75; 124pp; English.
XX
XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
XX AAB66932-AAB66941). The GLUTX proteins are related to the facultative
XX glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
XX function. The GLUTX proteins may be used in the diagnosis, prevention and
XX treatment of hexose transport disorders such as ischaemia, diabetes,
XX hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
XX neurodegenerative disease. The present sequence is a consensus sequence
XX for GLUTX1
XX
XX Sequence 478 AA;
XX
Query Match 90.5%; Score 2224.5; DB 4; Length 478;
Best Local Similarity 88.9%; Pred. No. 1.9e-204;
Matches 425; Conservative 23; Mismatches 29; Indels 1; Gaps 1;

QY 1 MTPEDPEETQPLLPGGGAPRRRVFLAAFAAALGFLGFGFALGYSSPAIPSLQRAAPP 60
DB 1 MSPEDEQETQLRPPGAPRRRVFLAAFAAALGFLGFGFALGYSSPAIPSLRTAPP 60
QY 61 APRLDDAASWFGAVVTLGAAAGVVLGWLVDRAKRLSLLCVSPFVAGFAVITAQDV 120
DB 61 ALRLGDAASWFGAVVTLGAAAGVVLGWLVDRAKRLSLLCVSPFVAGFAVITAQDV 120
QY 121 WMLLGGRLTGLACGVASLIVAPYIIEIAYPAVRGLGSCVQLMVVVVGLLACGVLE 180
DB 121 WMLLGGRLTGLACGVASLIVAPYIIEIAYPAVRGLGSCVQLMVVVVGLLACGVLE 180
QY 181 WRWLAVLGCVPVPSMLLMLCMFETPRFLITQHRROEAMAAALFLWGSQGWEDDPPIGAE 240
DB 181 WRWLAVLGCVPVPTMLLMLCMFETPRFLITQHQYQEAAMAAALFLWGSQGWEDDPVGA 240
QY 241 -QSFHLALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVNFYAEITFEAKFKDSSLASVVG 299

```

```

DB 241 HQGFQALLRRPGIYKPLIIGISLMWFQQLSGVNAVNFYAEITFEAKFKDSSLASVVG 300
QY 300 VIQVLTFAVAALIMDRAGRELLLVLSGVNVVFTSARFAYFKLTQSGPGNSHVAISAPV 359
DB 301 IIQVLTFAVAALIMDRAGRELLLVLSGVNVVFTSARFAYFKLTQSGPGNSHVAISAPV 360
QY 360 SAQPVDSVGLAVGSMCLFIAGPAVGWGPFWLLMSEIFPLHVKGVATGICVLTNWL 419
DB 361 SAEPDVSVGLAVGSMCLFIAGPAVGWGPFWLLMSEIFPLHVKGVATGICVLTNWF 420
QY 420 MAFLVTKFESLMEVLPYCAFWLASAFIPSVLFTLCVPEYKTKLEQITAHFEGR 477
DB 421 MAFLVTKFESLMEVLPYCAFWLASAFIPSVLFTLCVPEYKTKLEQITAHFEGR 478

RESULT 5
AAB66933
ID AAB66933 standard; protein; 478 AA.
XX
XX AC AAB66933;
XX
XX DT 17-APR-2001 (first entry)
XX
XX DE Rat GLUTX1.
XX
XX KW Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;
XX hexose transport disorder; ischaemia; diabetes; hyperglycaemia;
XX hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.
XX
XX OS Rattus sp.
XX
XX PN WO200104145-A2.
XX
XX PD 18-JAN-2001.
XX
XX PF 14-JUL-2000; 2000WO-IB001042.
XX
XX PR 14-JUL-1999; 99US-0143907P.
XX
XX PR 27-AUG-1999; 99US-0151140P.
XX
XX PR 23-FEB-2000; 2000US-0184285P.
XX
XX PR 13-JUL-2000; 2000US-00616132.
XX
XX PA (UYLA-) UNIV LAUSANNE.
XX
XX PI Thorens B, Ibberson M, Uldry M;
XX
XX WPI; 2001-112615/12.
XX
XX DR N-PSDB; AAF55866.
XX
XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
XX prevention, diagnosis and treatment of hexose transport disorders, e.g.
XX ischemia and diabetes.
XX
XX Claim 11; Page 71-72; 124pp; English.
XX
XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
XX AAB66932-AAB66941). The GLUTX proteins are related to the facultative
XX glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
XX function. The GLUTX proteins may be used in the diagnosis, prevention and
XX treatment of hexose transport disorders such as ischaemia, diabetes,
XX hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
XX neurodegenerative disease. The present sequence is rat GLUTX1
XX
XX Sequence 478 AA;
XX
Query Match 88.0%; Score 2162.5; DB 4; Length 478;
Best Local Similarity 85.4%; Pred. No. 1.7e-198;
Matches 408; Conservative 33; Mismatches 36; Indels 1; Gaps 1;

QY 1 MTPEDPEETQPLLPGGGAPRRRVFLAAFAAALGFLGFGFALGYSSPAIPSLQRAAPP 60
DB 1 MSPEDEQETQLRPPGAPRRRVFLAAFAAALGFLGFGFALGYSSPAIPSLRTAPP 60

```


QY 121 WMLGRLTGLACGVASLVPVYISETAYPAVRGLLSCVQLMVVVGILLAYLAGWVLE 180
Db 74 -----VYISEIAYPAVRGLLSCVQLMVVVGILLAYLAGWVLE 111
QY 181 WRWLAVLGVCPSPSLMLLLCMFMPETPRFLLTOHRRQEAAMALRFLNGSQGHEDDPGAE 240
Db 112 WRWLAVLGVCPSPSLMLLLCMFMPETPRFLLTOHRRQEAAMALRFLNGSQGHEDDPGAE 171
QY 241 QSFHLALLRQPGIKYKPFIIIGVSLMAFQQLSGVNAVVFYAEITFEAKFKDSSLASVVVG 300
Db 172 -----QQLSGVNAVVFYAEITFEAKFKDSSLASVVVG 205
QY 301 IQVLTFAVALINDRAGRRLLLVSLGVWVFSTSAFGAYFKLTQGGPGNSHVAISAPVS 360
Db 206 IQVLTFAVALINDRAGRRLLLVSLGVWVFSTSAFGAYFKLTQGGPGNSHVAISAPVS 265
QY 361 AQPVDASVGLAVLAVGSMCLFIAGFVAGWGPIPWLLMSIFPLHVKGVATGICVLTNWL 420
Db 266 AQPVDASVGLAVLAVGSMCLFIAGFVAGWGPIPWLLMSIFPLHVKGVATGICVLTNWL 325
QY 421 AFLVTKFSSLMELVLRPYGAFWLASAFICPSVLTFLFCVPETKGTLEQITAHFEGR 477
Db 326 AFLVTKFSSLMELVLRPYGAFWLASAFICPSVLTFLFCVPETKGTLEQITAHFEGR 382
RESULT 9
ADL33342
ID ADL33342 standard; protein; 353 AA.
AC ADL33342;
XX
DT 20-MAY-2004; (first entry)
XX
DE Human transporter and ion channel (TRICH) protein #46.
XX
KW anti-HIV; antiallergic; antiinflammatory; antianemic; antiparkinsonian;
KW nootropic; anticonvulsant; antiarteriosclerotic; antiaesthetic;
KW immunosuppressive; antithyroid; cytostatic; hepatotropic; dermatological;
KW antidiabetic; aephrotropic; antigout; thyromimetic; neuroprotective;
KW osteopathic; antiarthritic; antiparasitic; antihelminthic; antiparasitic;
KW uropathic; ophthalmological; antirheumatic; hemostatic; antibacterial;
KW virucide; phototoxide; fungicide; gene therapy.
XX
OS Homo sapiens.
XX
XX WO2003083085-A2.
XX
PD 09-OCT-2003.
XX
PF 27-MAR-2003; 2003WO-US0009797.
XX
PR 28-MAR-2002; 2002US-0368840P.
PR 26-APR-2002; 2002US-0375637P.
XX
PA (INCY-) INCYTE CORP.
XX
PI Marquis JP, Lee SY, Emerling BM, Hafalia AJA, Khare R, Kable AE;
PI Richardson TW, Swarnakar A, Chawla NK, Becha SD, Mason PM;
PI Elliott VS, Ramkumar J, Griffin JA, Tran UK, Ison CH, Lindquist EA;
PI Jiang X, Jackson AA, Wilson AD, Jin P, Chang H;
XX
DR WPI; 2003-833535/77.
DR N-PSDB; ADL33401.
XX
XX New human transporters and ion channels (TRICH) and polynucleotides,
PT useful for diagnosing, treating or preventing autoimmune or inflammatory
PT disorders (e.g. AIDS, allergy or anemia), multiple sclerosis, cancer or
PT hepatitis.
XX
XX Claim 1; SEQ ID NO 46; 405pp; English.
PS
XX The invention relates to an isolated polypeptide (I), which is a human
CC intracellular signaling molecule, which is a human intracellular

CC signaling molecule, a naturally occurring amino acid sequence at least 90%
CC -96% identical to it or a biologically active fragment or an immunogenic
CC fragment of the polypeptide. The human TRICH, polynucleotides, agonists
CC and antagonists are useful for diagnosing, treating or preventing
CC disorders associated with aberrant expression of TRICH, particularly cell
CC proliferative disorders (e.g. arteriosclerosis, atherosclerosis,
CC cirrhosis, hepatitis, paroxysmal nocturnal hemoglobinuria, polycythemia
CC vera, psoriasis, primary thrombocytopenia or cancer), developmental
CC disorders (e.g. renal tubular acidosis, anemia or mental retardation),
CC neurological disorders (e.g. Alzheimer's disease, Parkinson's disease or
CC epilepsy), autoimmune/inflammatory disorders (e.g. AIDS, allergies,
CC asthma, autoimmune thyroiditis, contact dermatitis, Crohn's disease,
CC diabetes mellitus, glomerulonephritis, Goodpasture's syndrome, gout,
CC Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,
CC multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
CC syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,
CC bacterial, fungal, parasitic, protozoan or helminthic infections. The
CC polynucleotides encoding TRICH are useful for creating transgenic animals
CC to model human disease. This sequence corresponds to one of the proteins
CC of the invention.
XX
SQ Sequence 353 AA;
Query Match 67.7%; Score 1664; DB 7; Length 353;
Best Local Similarity 100.0%; Pred. No. 8e-151;
Matches 326; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MTPDEPTEOTPLLPGGSGAPRGRRVFLAFAAALGPLSGFALGVSSPAISLQRAAPP 60
Db 1 MTPDEPTEOTPLLPGGSGAPRGRRVFLAFAAALGPLSGFALGVSSPAISLQRAAPP 60
QY 61 APRLDDAASWFGAVVTGAAAGGVILGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDV 120
Db 61 APRLDDAASWFGAVVTGAAAGGVILGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDV 120
QY 121 WMLGRLTGLACGVASLVPVYISETAYPAVRGLLSCVQLMVVVGILLAYLAGWVLE 180
Db 121 WMLGRLTGLACGVASLVPVYISETAYPAVRGLLSCVQLMVVVGILLAYLAGWVLE 180
QY 181 WRWLAVLGVCPSPSLMLLLCMFMPETPRFLLTOHRRQEAAMALRFLNGSQGHEDDPGAE 240
Db 181 WRWLAVLGVCPSPSLMLLLCMFMPETPRFLLTOHRRQEAAMALRFLNGSQGHEDDPGAE 240
QY 241 QSFHLALLRQPGIKYKPFIIIGVSLMAFQQLSGVNAVVFYAEITFEAKFKDSSLASVVVG 300
Db 241 QSFHLALLRQPGIKYKPFIIIGVSLMAFQQLSGVNAVVFYAEITFEAKFKDSSLASVVVG 300
QY 301 IQVLTFAVALINDRAGRRLLLVLSG 326
Db 301 IQVLTFAVALINDRAGRRLLLVLSG 326
RESULT 10
ABB9717
ID ABB9717 standard; protein; 326 AA.
XX
AC ABB9717;
XX
DT 24-MAY-2002 (first entry)
XX
DE Human polypeptide SEQ ID NO 2093.
XX
KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;
KW antiallergic; hepatotropic; antidiabetic; antiinflammatory; antituber;
KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;
KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;
KW neurological disease; infection; human; secreted protein.
XX
OS Homo sapiens.
XX
XX WO200190304-A2.
XX
XX 29-NOV-2001.

```
XX 18-MAY-2001; 2001WO-US016450.
XX
XX 19-MAY-2000; 2000US-0205515P.
XX
XX (HUMA-) HUMAN GENOME SCI INC.
XX
XX Birse CE, Rosen CA;
XX
XX WPI; 2002-122018/16.
XX
XX N-PSDB; ABL90126.
XX
XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
XX prevention of neural, immune system, muscular, reproductive,
XX gastrointestinal, pulmonary, cardiovascular, renal and proliferative
XX disorders.
XX
XX Claim 11; SEQ ID NO 2093; 2081pp + Sequence Listing; English.
XX
XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
XX (AB89040-AB89044) useful for preventing, treating or ameliorating
XX medical conditions e.g. by protein or gene therapy. The genes are
XX isolated from a range of human tissues disclosed in the specification.
XX The nucleic acids, proteins, antibodies and (ant)agonists are useful in
XX the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
XX ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
XX breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
XX disorders e.g. Addison's disease, allergies, autoimmune haemolytic
XX anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
XX multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
XX cardiovascular disorders such as myocardial ischaemias; (d) wound healing
XX ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
XX infectious diseases such as viral, bacterial, fungal and parasitic
XX infections. Note: The sequence data for this patent did not form part of
XX the printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
XX Sequence 326 AA;
XX
XX Query Match 60.7%; Score 1491; DB 5; Length 326;
XX Best Local Similarity 96.3%; Pred. No. 2.9e-134;
XX Matches 289; Conservative 1; Mismatches 10; Indels 0; Gaps 0;
XX
XX 164 MVVVGILLAYLAGVLEWLVAVLGCVPSSLMLLMCFMPEPRFLTTOHRRQEAVALR 223
XX |
XX 1 MVVVGILLAYLAGVLEWLVAVLGCVPSSLMLLMCFMPEPRFLTTOHRRQEAVALR 60
XX |
XX 224 FLWGSQGWEDPPIGAEQSFHLALRQPGIYKPFIIIGVSLMAFQQLSGVNAVWYAEITF 283
XX |
XX 61 FLWGSQGWEDPPIGAEQSFHLALRQPGIYKPFIIIGVSLMAFQQLSGVNAVWYAEITF 120
XX |
XX 284 BEAKFKDSSLASVVVGIVQVLFVAVAAALMDRGRRLLLVLSGVVWVFTSAFGYFKLT 343
XX |
XX 121 BEAKFKDSSLASVVVGIVQVLFVAVAAALMDRGRRLLLVLSGVVWVFTSAFGYFKLT 180
XX |
XX 344 QGPGNSSHVAISAPVSAPVDASVGLAVGSMCLFTAGFAGVWGPPIWLLMSEIRPL 403
XX |
XX 181 QGPGNSSHVAISAPVSAPVDASVGLAVGSMCLFTAGFAGVWGPPIWLLMSEIRPL 240
XX |
XX 404 HVKGVATGICVLTNWLMALFVTKFEFSLMEVLRYPGAFWLASAFICFSVLFVFCVPEPK 463
XX |
XX 241 HVKGVATGICVLTNWLMALFVTKFEFSLMEVLRYPGAFWLASAFICFSVLFVFCVPEPK 300
XX |
XX
XX RESULT 11
XX AA06579
XX ID AA06579 standard; protein; 262 AA.
XX
XX AC AA06579;
XX
XX XX 25-SEP-2001 (first entry)
XX
XX Human protein having hydrophobic domain, HP10784.
```

```
XX Human; hydrophobic domain; gene therapy; nutritional supplement;
XX cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
XX multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
XX haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
XX Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
XX haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
XX contraceptive; antiinfertility; antiinflammatory.
XX
XX Homo sapiens.
XX
XX WO200149728-A2.
XX
XX 12-JUL-2001.
XX
XX 28-DEC-2000; 2000WO-JP009359.
XX
XX 06-JAN-2000; 2000JP-0000585.
XX
XX 06-JAN-2000; 2000JP-0000588.
XX
XX 11-JAN-2000; 2000JP-00002299.
XX
XX 03-FEB-2000; 2000JP-00026862.
XX
XX 03-MAR-2000; 2000JP-00058367.
XX
XX (PROT-) PROTEGENE INC.
XX (SAGA ) SAGAMI CHEM RES CENT.
XX
XX Kato S, Kimura T;
XX
XX WPI; 2001-418355/44.
XX N-PSDB; AAD12574.
XX
XX Human proteins with hydrophobic domains and the nucleic acids encoding
XX them, useful for preventing diagnosing and treating e.g. cancer,
XX Alzheimer's and inflammation.
XX
XX Claim 1; Page 75; 563pp; English.
XX
XX The present sequence is human protein with hydrophobic domain, HP10784.
XX The polynucleotide and polypeptide of the invention may be used in the
XX prevention, diagnosis and treatment of diseases associated with
XX inappropriate polypeptide expression. The polynucleotides may be used to
XX produce the polypeptide, by inserting the nucleic acids into a host cell
XX and culturing the cell to express the protein. The polynucleotides and
XX its complementary sequences may also be used as DNA probes in diagnostic
XX assays and also used in gene therapy. The polypeptides may also be used
XX as antigens in the production of antibodies and in assays to identify
XX modulators of polypeptide expression and activity. The polypeptides and
XX nucleic acids may be used as nutritional supplements, to modulate
XX cytokine and cell proliferation activity, to modulate immune stimulation
XX or suppression (e.g. for the treatment of microbial infections and
XX autoimmune disorders such as multiple sclerosis, rheumatoid arthritis and
XX insulin-dependent diabetes), to modulate haematopoiesis, to modulate
XX tissue growth activity (e.g. for the treatment of Parkinson's disease,
XX Huntington's disease and Alzheimer's disease), to modulate activin and
XX inhibin activity (e.g. for controlling fertility), to modulate
XX chemotactic and chemokinetic activity, to modulate haemostatic and
XX thrombolytic activity, to modulate receptor ligand activity, to modulate
XX inflammation and to inhibit tumour growth
XX
XX Sequence 262 AA;
XX
XX Query Match 46.2%; Score 1135; DB 4; Length 262;
XX Best Local Similarity 91.4%; Pred. No. 3.3e-100;
XX Matches 223; Conservative 2; Mismatches 7; Indels 12; Gaps 1;
XX
XX 1 MTPEDPETQPLGPPGGSAPGRVRVFLAAFAAALPLSFALGYSFALPSLQRAAPP 60
XX |
XX 1 MTPEDPETQPLGPPGGSAPGRVRVFLAAFAAALPLSFALGYSFALPSLQRAAPP 60
XX |
XX 61 APRLDDAAASWFGAVVTGLAAAGGVLGWLVDRAGRKLILLICSVFVAGFAVITAAQDV 120
XX |
XX 61 APRLDDAAASWFGAVVTGLAAAGGVLGWLVDRAGRKLILLICSVFVAGFAVITAAQDV 120
XX |
XX
```

Qy	121	WMLLGGRLTGTGACGVASIVAPVYISETAYPAVPGLLGSCVQLMVMVVGITLLAVLAGWVLE	180
Db	121	WMLLGGRLTGTGACGVASIVAPVYISETAYPAVPGLLGSCVQLMVMVVGITLLAVLAGWVLE	180
Qy	181	WRWLAVLGCVPPSLLMLLLMCFMPTPRLFTQHRROBAM	228
Db	181	WRWLAVLGCVPPSLLMLLLMCFMPTPRLFTQHRROBAAFGLVRCGHGVQHECLRLLQA	240
Qy	229	EQGW	232
Db	241	DPGW	244

RESULT 12	
ABU11283	
ID	ABU11283 standard; protein; 248 AA.
XX	
XX	ABU11283;
XX	
XX	
DT	10-FEB-2003 (first entry)
XX	
XX	cDNA encoding human cancer suppressing protein PF7425.
DE	
XX	
XX	Human; cancer suppressing protein; cancer.
KW	
XX	
XX	
OS	Homo sapiens.
XX	
PN	CN1351081-A.
XX	
XX	
PD	29-MAY-2002.
XX	
PF	31-OCT-2000; 2000CN-00127102.
XX	
XX	
PR	31-OCT-2000; 2000CN-00127102.
XX	
XX	
PA	(SHAN-) SHANGHAI INST ONCOLOGY.

Db	181	QGGFGNSSHVAISAPVDAQPVDSVGLAWLAVGSMCLFIAG-----GPOALMSLLACLRF	2335
Qy	403	LHVK 406	
Db	236	LHLQ 239	
RESULT 13			
AAB66938			
ID	AAB66938 standard; protein; 503 AA.		
XX	AC	AAB66938;	
XX	AC		
DT	17-APR-2001 (first entry)		
XX			
DE	Rat GLUTX3.		
XX			
KW	Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;		
KW	hexose transport disorder; ischaemia; diabetes; hyperglycaemia;		
KW	hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.		
XX			
OS	Rattus sp.		

	New human protein with cancer cell growth suppressing function and a polynucleotide encoding it, for treating diseases, such as, cancer.
PT	
XX	
XX	
PS	Claim 1; Page 23 (disclosure); 39pp; Chinese.
XX	
CC	This invention relates to the cDNA and protein sequences of a novel human protein with cancer suppressing function. The invention also comprises a method for preparing the polypeptide by recombination, and an application of the polypeptide in treating diseases such as cancer, etc. Also disclosed in an antagonist of the polypeptide and its medical action. The present sequence represents a cancer suppressing protein of the invention
CC	
XX	
SQ	Sequence 248 AA;
	Query Match 46.0%; Score 1131; DB 5; Length 248;
	Best Local Similarity 93.0%; Pred. No. 7.4e-100;
	Matches 227; Conservative 5; Mismatches 6; Indels 6; Gaps 2
Qy	164 MVVGGILLAYLAGWLEWRWLAVLGCVPSPSLMLLMCMFMPETPRFLLTQHRREANMAALR 223
Db	- 1 MVVGGILLAYLAGWLEWRWLAVLGCVPSPSLMLLMCMFMPETPRFLLTQHRREANMAALR 60
Qy	224 FLWGSEQGWEDPPIGAEQSFHALLRQPGIYKYPFIIGVSLMAFOQLSGVNVMFYAETTF 283
Db	61 FLWGSEQGWEDPPIGAEQSFHALLRQPGIYKYPFIIGVSLMAFOQLSGVNVMFYAETTF 120
Qy	284 ERAKPKDSSLASVVGVGLQVLFTAAVALIMDRAGRLLLVLSGVVMVFSTSAFGAFVKLT 343
Db	121 ERAKPKDSSLASVVGVGLQVLFTAAVALIMDRAGRLLLVLSGVVMVFSTSAFGAFVKLT 180
Qy	344 QGGPGNNSHVAISAPVSQAQPVDASVGLAWLAVGSMCLFIAGFAVGWGP-IPMILLMSIEIF 402

QY	183	WLAVLGCVPPSLMLLLMCMFETPRFLLTQHRROEAMAAALRFLW--GSEQGWEDPPIG-	238
Db	194	WLAVAGEGVLVWMLLSFMPNSPRFLLSKSRDEALQAL--IWLADSEVHMEFEQIQD	251
QY	239	--AEQSFHALLR--OPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEITFEAK-FKDSSL	293
Db	252	NVRQSSRSVMAEAEPRVIRPILITVLMRFLOQLTGITPILVYLQTFDSTSVLPSQQ	311
QY	294	ASVVGVIOVLTAAALIMDRAGRLLLVSGVVMVFTSAFGAYFKLTQG--GPGNSS	351
Db	312	DAAIAGVAVKLLSVLAAVTMDLAGKRVLLYSASIMFVANLTGLYQLVPTLTNSIV	371
QY	352	HVAISAPVSAQVDASVGLAVGSMCLFTAGFVAGWGPIPWLLMSIFPIHVKVATG	411
Db	372	EIVTLGGTEQPPAAAFNYLTLLPLATMLFMGYANGWGPIITWLLMSEVLPLRARGVAS	431
QY	412	ICVLNWLMAFLVTFEFSLSMEVLRYPGAFWLASAFICFSLVLTLCVCPETKGTLEQIT	471
Db	432	LCVLSVLTAFVLTXYFLAVNAVAGLQVPPFFPSAICLLSLLTGCCTVETGRSLEQIE	491
QY	472	AHFEGR	477
Db	492	AFHTR	497
RESULT 14			
AA	66941	standard; protein; 507 AA.	
ID	AA66941		
AC	AA66941;		
DT	17-APR-2001	(first entry)	
DE	GLUTX3 consensus sequence.		
XX	GLUTX; gene therapy; vaccine; hexose transport modulator; human; rat;		
KW	hexose transport disorder; ischaemia; diabetes; hyperglycaemia;		
KW	hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.		
OS	Homo sapiens.		
OS	Rattus sp.		
XX	WO200104145-A2.		
XX	18-JAN-2001.		
XX	14-JUL-2000; 2000WO-IB001042.		
XX	14-JUL-1999; 99US-0143907P.		
PR	27-AUG-1999; 99US-0151140P.		
PR	23-FEB-2000; 2000US-0184285P.		
PR	13-JUL-2000; 2000US-00616132.		
XX	(UYLA-) UNIV. LAUSANNE.		
XX	Thorens B, Ibberson M, Uldry M;		
XX	WPI; 2001-112615/12.		
XX	Nucleic acids encoding GLUTX glucose transporter proteins, useful in the		
PT	prevention, diagnosis and treatment of hexose transport disorders, e.g.		
PT	ischemia and diabetes.		
XX	Claim 11; Page 83-84; 124pp; English.		
XX	The present invention relates to GLUTX proteins (AAF55865-AAF55871 and		
CC	AA66932-AA66941). The GLUTX proteins are related to the facultative		
CC	glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport		
CC	function. The GLUTX proteins may be used in the diagnosis, prevention and		
CC	treatment of hexose transport disorders such as ischaemia, diabetes,		
CC	hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a		
CC	neurodegenerative disease. The present sequence is a consensus sequence		

CC for GLUTX3

XX Sequence 507 AA;

QY Query Match 38.8%; Score 953; DB 4; Length 507;

Db Best Local Similarity 42.9%; Pred. No. 2.4e-82;

Matches 217; Conservative 78; Mismatches 167; Indels 44; Gaps 11;

QY 10 QPILG-----PPGSAAPRGR-----RVFLAAPAAALGPLSGFALCYSGPA 50

Db 3 EPLUGAEGPDYTFPEKPPASPGDRARVGTQNKRVFLATFAVLGNFSGFALVYVTSV 62

QY 51 IPSLQRAAPAPRLDDAAASWFGAVVTLCGAAGGVGLGWLVDRAGRKLSLLCSVPFVAG 110

Db 63 IPALERSLPDLHLTKSQASWFGSVFTLGAAGGLSAMTINDLLGKLSIMFSAVPSAAG 122

QY 111 FAVITAAQVMMLLGGRLLTGLACGVASLVAPVYIYSEIYPAVARGLLGSCVLMVVVGL 170

Db 123 YAVMAGAGLWMLLGLTGTGAGGLTAACTPVYVSEIAPPVGRGALGATPQLMAVFGSL 182

QY 171 LAYLAGVLEWRLAVLGVCPVPSLMLLMCFMPETPRFLLTQHRROEAMAAALRFLWSEQ 230

Db 183 SLVALGULLPWRWLAVAGKAPVLMILLISFPNSPRFLLSRGRDEEALRALAWLRGTDV 242

QY 231 G--WEDPPIG--AEQSFHL--ALLRQPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEITF 283

Db 243 DVHWEFGIQDNVRRQSSRSVMAEAPRHVCRPITVALLMRLIQLTGITPILVYLQSF 302

QY 284 EBAKF-----KSSIASVVGVIOVLTAAALIMDRAGRLLLVSGVVMVFTSAFGAY 339

Db 303 DSTAVLLPPKDD--AAIVGAVRLLSVLTAALTMDLAGRVLLFVSAATMFAANLTGLY 359

QY 340 FKLTQGGFGNSSHVAISAPVS-----AQPVDASVG--LAWLAVGSMCLFTIAGFVAGWGPI 392

Db 360 IHF---GPRLSNPNSTAGLESWSGDLAQPLAAPAGVLTLPVLLATMLFMGYANGWGPI 416

QY 393 PWLLMSIFPIHVKVATGICVLTNWLMAFLVTKFSSLSMEVLRYPGAFWLASAFICFSV 452

Db 417 TWLLMSEVLPLRARGVASGLCVLASMLTAFVLTKSLFPVVSTFGLOVPPFLFPAICLVSL 476

QY 453 LFTLFCVCPETKGTLEQITAHFE-GR 477

Db 477 VFTGCCVETPKGRSLEQIESFFRTGR 502

RESULT 15

AA66937

ID AA66937 standard; protein; 507 AA.

AC AA66937;

DT 17-APR-2001 (first entry)

DE Human GLUTX3.

XX Human; GLUTX; gene therapy; vaccine; hexose transport modulator;

KW hexose transport disorder; ischaemia; diabetes; hyperglycaemia;

KW hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.

OS Homo sapiens.

XX WO200104145-A2.

XX 18-JAN-2001.

XX 14-JUL-2000; 2000WO-IB001042.

XX 14-JUL-1999; 99US-0143907P.

PR 27-AUG-1999; 99US-0151140P.

PR 23-FEB-2000; 2000US-0184285P.

PR 13-JUL-2000; 2000US-00616132.

XX (UYLA-) UNIV LAUSANNE.

```
XX Thorens B, Ibberson M, Uldry M;
PI WPI; 2001-112615/12.
DR N-PSDB; AAF55870.
XX
PT Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
PT ischemia and diabetes.
XX
PS Claim 11; Page 81-82; 124pp; English.
XX
CC The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
CC AAB6932-AAB6941). The GLUTX proteins are related to the facultative
CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
CC function. The GLUTX proteins may be used in the diagnosis, prevention and
CC treatment of hexose transport disorders such as ischaemia, diabetes,
CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
CC neurodegenerative disease. The present sequence is human GLUTX3
XX
SQ Sequence 507 AA;
Query Match : 38.6%; Score 948; DB 4; Length 507;
Best Local Similarity 42.3%; Pred. No. 7.1e-82;
Matches 217; Conservative 78; Mismatches 167; Indels 44; Gaps 11;
QY 10 QPLLG-----PPGGSAPGR-----RVFLAAFAAALGLSPFGALGYSSPA 50
Db 3 EPLLAGEPDYDTFFPEKPPSPQDRARVGTQLQNKRVFLATFAAVLGNFSGYALVYTSVP 62
QY 51 IPSLQRAAPPAPRLDAAASWFGAVVTLGAAGVGLGGWLVDRAGRKLSLLCSVEFVAG 110
Db 63 IPALERSLPDLHLTKSQASWFGSVFTLGAAGGLSAMLNDLLGRKLSIMFSAPVSAAG 122
QY 111 FAVITAAQDVMMLLGRLTLGACVASLVAPVYISEIAYPAVRGLLSCVOLWVVGIL 170
Db 123 YAVMAGHGLMMLLLGRTLTGAGGLTAACIPVYVSEIAPPVGRGALGATPQLMAVFGSL 182
QY 171 LAYLAGVLEWRMLAVLGCVPPLMLLLMCFMPEPRFLITQHRROEAMALRFLMGSEQ 230
Db 183 SLVALGULLPWRWLAVAGAPVLMILLISFMPNSPRFLLSRGRDEALRALAWLGTDV 242
QY 231 G--WEDPIIG--AEQSFHL--ALLRQPGIYKFFIIGVSLMAFQQLSGVNAVNFYAEITF 283
Db 243 DVHWEFGQIQDNVRRQSSRVSAEARAPHVCRPTVALLMRLIQLTGITPILVYLQSF 302
QY 284 BEAKP----KDSLSASVVYQVIOVFTAVAALIMDRGRLLLVLSGVVWVFSTSAFGAY 339
Db 303 DSTAVLLPKDD--AAIVGXVRLLSVLTAALTMDLAGRKVLLFVSAAIMFAANLTIGLY 359
QY 340 FKLTQGGPGNSHVAISAPVS-----AOPVDASVG-LAWLAVGSMCLFIAGFVAGWGPI 392
Db 360 IHF---GPRPLSNSTAGLESWSGDLAQPLAAPAGYLTIVPLLATMLFIMGYAVGWGPI 416
QY 393 PMLMSEIIPPLHKVGVATGICVLTNWMAFLVTKFESSLMELRPHYGAFWLASAFCIFSV 452
Db 417 TWLLMSEVLPLRARGVAGLCVLASMLTAFVLTKSFLPVVSTFGLQVPFLFAAICLVSL 476
QY 453 LFTLFCVPETKGTLEQITAHFE-GR 477
Db 477 VFTGCCVPETKGRSLEQIESFFRTGR 502
```

Search completed: February 24, 2005, 05:49:51
Job time : 175 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 05:47:06 ; Search time 132 Seconds
(without alignments)
1182.530 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457

Sequence: 1 MTPEDPEETQPLGPPGSA.....CVPETKGTLEQTAHFEGR 477

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1380268 seqs, 327241040 residues

Total number of hits satisfying chosen parameters: 1380268

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/prodata/2/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/prodata/2/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/prodata/2/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/prodata/2/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/prodata/2/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/prodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/prodata/2/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/prodata/2/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 10: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 11: /cgn2_6/prodata/2/pubpaa/US09_PUBCOMB.pep.*
- 12: /cgn2_6/prodata/2/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/prodata/2/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/prodata/2/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/prodata/2/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/prodata/2/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/prodata/2/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/prodata/2/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/prodata/2/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/prodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2457	100.0	477	14	US-10-168-651-1
2	2450	99.7	477	9	US-09-886-954-1
3	1491	60.7	326	15	US-10-264-237-2093
4	1135	46.2	262	15	US-10-169-395-9
5	948	38.6	507	14	US-10-157-031-359
6	948	38.6	507	16	US-10-755-889-110
7	723	29.4	501	13	US-10-051-909-38
8	716	29.1	484	15	US-10-424-599-275264
9	702	28.6	501	16	US-10-437-963-163544
10	702	28.6	523	15	US-10-425-114-63431
11	701.5	28.6	509	15	US-10-425-114-59933
12	658.5	26.8	553	13	US-10-051-909-36
13	639	26.0	481	15	US-10-310-154-725

14	615	25.0	488	10	US-09-794-822-11	Sequence 11, Appl
15	615	25.0	488	13	US-10-094-059-4	Sequence 4, Appl
16	615	25.0	488	14	US-10-170-528-5	Sequence 5, Appl
17	615	25.0	488	14	US-10-162-012-46	Sequence 46, Appl
18	615	25.0	488	14	US-10-062-960B-4	Sequence 4, Appl
19	615	25.0	488	14	US-10-144-624-4	Sequence 4, Appl
20	615	25.0	488	15	US-10-162-102-46	Sequence 46, Appl
21	615	25.0	488	15	US-10-038-854-170	Sequence 170, Appl
22	613.5	25.0	487	9	US-09-795-693-27	Sequence 27, Appl
23	613.5	25.0	487	13	US-10-095-139-14	Sequence 14, Appl
24	613.5	25.0	487	14	US-10-156-239-27	Sequence 27, Appl
25	613.5	25.0	487	14	US-10-199-485-27	Sequence 27, Appl
26	613.5	25.0	487	15	US-10-391-399-45	Sequence 45, Appl
27	609.5	24.8	502	16	US-10-437-963-166774	Sequence 166774, A
28	607	24.7	461	15	US-10-369-493-23371	Sequence 23371, A
29	604.5	24.6	592	15	US-10-425-114-63791	Sequence 63791, A
30	604.5	24.6	596	16	US-10-437-963-177463	Sequence 177463, A
31	604.5	24.6	601	15	US-10-425-114-63432	Sequence 63432, A
32	603	24.5	472	13	US-10-024-623-29	Sequence 29, Appl
33	603	24.5	472	14	US-10-154-419-79	Sequence 79, Appl
34	603	24.5	472	14	US-10-146-733-74	Sequence 74, Appl
35	603	24.5	472	15	US-10-369-493-23536	Sequence 23536, A
36	598.5	24.4	465	15	US-10-282-122A-56366	Sequence 56366, A
37	598	24.3	486	9	US-09-860-232A-7	Sequence 7, Appl
38	593.5	24.2	525	15	US-10-425-114-63426	Sequence 63426, A
39	591.5	24.1	464	13	US-10-024-623-28	Sequence 28, Appl
40	591.5	24.1	464	14	US-10-154-419-78	Sequence 78, Appl
41	591.5	24.1	464	14	US-10-146-733-73	Sequence 73, Appl
42	591.5	24.1	464	15	US-10-369-493-23552	Sequence 23552, A
43	591.5	24.1	464	15	US-10-282-122A-42918	Sequence 42918, A
44	591.5	24.1	523	16	US-10-437-963-124519	Sequence 124519, A
45	588.5	24.0	535	9	US-09-795-693-20	Sequence 20, Appl

ALIGNMENTS

RESULT 1
US-10-168-651-1
Sequence 1, Application US/10168651
Publication No. US20030171275A1
GENERAL INFORMATION:
APPLICANT: INCYTE GENOMICS, INC.
APPLICANT: BAUGHN, Mariah R.
APPLICANT: BURFORD, Neil
APPLICANT: AU-YOUNG, Janice
APPLICANT: LU, Dyung Aina M.
APPLICANT: YANG, Junming
APPLICANT: REDDY, Roopa
APPLICANT: LAL, Preeti
APPLICANT: HILLMAN, Jennifer L.
APPLICANT: AZIMZAI, Valda
APPLICANT: YUE, Henry
APPLICANT: NGUYEN, Dannel B.
APPLICANT: YAO, Monique G.
APPLICANT: GANDHI, Ameeni R.
APPLICANT: TANG, Y. Tom
APPLICANT: KHAN, Farrah A.
TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
FILE REFERENCE: PI-0005 PCT
CURRENT APPLICATION NUMBER: US/10/168,651
PRIOR FILING DATE: 2002-06-21
PRIOR APPLICATION NUMBER: 60/172,000; 60/176,083; 60/177,332; 60/178,572; 60/179,758
60/181,625
PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;
2000-02-10
NUMBER OF SEQ ID NOS: 54
SOFTWARE: PERL Program
SEQ ID NO 1
LENGTH: 477
TYPE: PRT
ORGANISM: Homo sapiens
FEATURE:

```

; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No. US20030171275A1 1416107CD1
US-10-168-651-1

Query Match      100.0%; Score 2457; DB 14; Length 477;
Best Local Similarity 100.0%; Pred. No. 2.9e-203;
Matches 477; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MTPDEETQPLLGPPGSGAPRRVFLAAFAAALGPLSFGFALGYSSPAIPSLQRAAPP 60
DB 1 MTPDEETQPLLGPPGSGAPRRVFLAAFAAALGPLSFGFALGYSSPAIPSLQRAAPP 60
QY 61 APRLDDAAASWFGAVVTLGAAAGGVLGGWLVDRAGKLSLLCSVPFVAGFAVITAADV 120
DB 61 APRLDDAAASWFGAVVTLGAAAGGVLGGWLVDRAGKLSLLCSVPFVAGFAVITAADV 120
QY 121 WMLLGRLLTGLACGVASLVAIVYSEIAYPAVRGLGSCVQLMVAIVYSEIAYLAGWVLE 180
DB 121 WMLLGRLLTGLACGVASLVAIVYSEIAYPAVRGLGSCVQLMVAIVYSEIAYLAGWVLE 180
QY 181 WRWLAVLGCVPVPSLMLLMCFMPEPRFLLTOHRRQEAAMALRFLWGSEQGWEDDPPICAE 240
DB 181 WRWLAVLGCVPVPSLMLLMCFMPEPRFLLTOHRRQEAAMALRFLWGSEQGWEDDPPICAE 240
QY 241 QSFHALLRQPGIYKPIIGVSLMAFQQLSGVNAVNFYAEITFEAKPKDSSLASVWVG 300
DB 241 QSFHALLRQPGIYKPIIGVSLMAFQQLSGVNAVNFYAEITFEAKPKDSSLASVWVG 300
QY 301 IOVLFTAVAAALINDRAGRLLLVLSGVNVVFTSFAFGYFKLTQGGPGNSSHVAISAPVS 360
DB 301 IOVLFTAVAAALINDRAGRLLLVLSGVNVVFTSFAFGYFKLTQGGPGNSSHVAISAPVS 360
QY 361 AQVDASVGLAWLAVGSMCLFIAGFAVGWGPPIWLLMSEIFPLHVKGVATGICVLTNWLM 420
DB 361 AQVDASVGLAWLAVGSMCLFIAGFAVGWGPPIWLLMSEIFPLHVKGVATGICVLTNWLM 420
QY 421 AFLVTKFSSLMELVLPYGAFWLASAFICFVSFLTFCVPETKGTLEQITAHFEGR 477
DB 421 AFLVTKFSSLMELVLPYGAFWLASAFICFVSFLTFCVPETKGTLEQITAHFEGR 477

RESULT 2
US-09-886-954-1
; Sequence 1, Application US/09886954
; Publication No. US20020038464A1
; GENERAL INFORMATION:
; APPLICANT: Chazron, Maureen J.
; APPLICANT: Katz, Ellen B.
; TITLE OF INVENTION: NOVEL GLUCOSE TRANSPORTER/SENSOR PROTEIN AND USES THEREOF
; FILE REFERENCE: 96700/667
; CURRENT APPLICATION NUMBER: US/09/886,954
; CURRENT FILING DATE: 2001-06-21
; NUMBER OF SEQ ID NOS: 1
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 1
; LENGTH: 477
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-886-954-1

Query Match      99.7%; Score 2450; DB 9; Length 477;
Best Local Similarity 99.8%; Pred. No. 1.2e-202;
Matches 476; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MTPDEETQPLLGPPGSGAPRRVFLAAFAAALGPLSFGFALGYSSPAIPSLQRAAPP 60
DB 1 MTPDEETQPLLGPPGSGAPRRVFLAAFAAALGPLSFGFALGYSSPAIPSLQRAAPP 60
QY 61 APRLDDAAASWFGAVVTLGAAAGGVLGGWLVDRAGKLSLLCSVPFVAGFAVITAADV 120
DB 61 APRLDDAAASWFGAVVTLGAAAGGVLGGWLVDRAGKLSLLCSVPFVAGFAVITAADV 120
QY 121 WMLLGRLLTGLACGVASLVAIVYSEIAYPAVRGLGSCVQLMVAIVYSEIAYLAGWVLE 180
DB 121 WMLLGRLLTGLACGVASLVAIVYSEIAYPAVRGLGSCVQLMVAIVYSEIAYLAGWVLE 180
QY 181 WRWLAVLGCVPVPSLMLLMCFMPEPRFLLTOHRRQEAAMALRFLWGSEQGWEDDPPICAE 240
DB 181 WRWLAVLGCVPVPSLMLLMCFMPEPRFLLTOHRRQEAAMALRFLWGSEQGWEDDPPICAE 240
QY 241 QSFHALLRQPGIYKPIIGVSLMAFQQLSGVNAVNFYAEITFEAKPKDSSLASVWVG 300
DB 241 QSFHALLRQPGIYKPIIGVSLMAFQQLSGVNAVNFYAEITFEAKPKDSSLASVWVG 300
QY 301 IOVLFTAVAAALINDRAGRLLLVLSGVNVVFTSFAFGYFKLTQGGPGNSSHVAISAPVS 360
DB 301 IOVLFTAVAAALINDRAGRLLLVLSGVNVVFTSFAFGYFKLTQGGPGNSSHVAISAPVS 360
QY 361 AQVDASVGLAWLAVGSMCLFIAGFAVGWGPPIWLLMSEIFPLHVKGVATGICVLTNWLM 420
DB 361 AQVDASVGLAWLAVGSMCLFIAGFAVGWGPPIWLLMSEIFPLHVKGVATGICVLTNWLM 420
QY 421 AFLVTKFSSLMELVLPYGAFWLASAFICFVSFLTFCVPETKGTLEQITAHFEGR 477
DB 421 AFLVTKFSSLMELVLPYGAFWLASAFICFVSFLTFCVPETKGTLEQITAHFEGR 477

US-10-264-237-2093
; Sequence 2093, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birse et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: PatentIn Ver. 3.1
; SEQ ID NO 2093
; LENGTH: 326
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: MISC FEATURE
; LOCATION: (76)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (87)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (249)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (253)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (258)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (261)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; FEATURE:
; NAME/KEY: MISC FEATURE
; LOCATION: (316)
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
; OTHER INFORMATION: Xaa equals any of the twenty naturally occurring L-amino acids
US-10-264-237-2093

```

Query Match 60.7%; Score 1491; DB 15; Length 326;
Best Local Similarity 96.3%; Pred. No. 3.8e-120;
Matches 289; Conservative 1; Mismatches 10; Indels 0; Gaps 0;

QY 164 MVVVGILLAYLAGWLEWRWLAIVGCVPSLSMLLMCFMPETPRFLLTQHRQEAALR 223
DB 1 MVVVGILLAYLAGWLEWRWLAIVGCVPSLSMLLMCFMPETPRFLLTQHRQEAALR 60

QY 224 FLWGSEQWEDPPIGAEQSFHLALRQPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEITF 283
DB 61 FLWGSEQWEDPPIGAEQSFHLALRQPGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEITF 120

QY 284 BEAKPKDSSLASVVVGVIVQLFTVAALIMDRAGRELLLVLSGVVNVFSTSAFGAYFKLT 343
DB 121 BEAKPKDSSLASVVVGVIVQLFTVAALIMDRAGRELLLVLSGVVNVFSTSAFGAYFKLT 180

QY 344 QGGPGNSSHVATSAFVSAQPVDSVGLAVGSMCLFIAGFVAGVGPFWLLMSIFPL 403
DB 181 QGGPGNSSHVATSAFVSAQPVDSVGLAVGSMCLFIAGFVAGVGPFWLLMSIFPL 240

QY 404 HVKGATGICVLNTWMLMAFLVTKFSSLMELVLPYGAFWLASAFCLFSVLTFCVPETK 463
DB 241 HVKGATGICVLNTWMLMAFLVTKFSSLMELVLPYGAFWLASAFCLFSVLTFCVPETK 300

RESULT 4
US-10-169-395-9
; Sequence 9, Application US/10169395
; Publication No. US20040034192A1
; GENERAL INFORMATION:
; APPLICANT: KATO, Seishi
; APPLICANT: KIMURA, Tomoko
; TITLE OF INVENTION: HUMAN PROTEINS HAVING HYDROPHOBIC DOMAINS AND DNAS ENCODING
; FILE REFERENCE: 01997.015100.US
; CURRENT APPLICATION NUMBER: US/10/169,395
; PRIOR FILING DATE: 2002-11-29
; PRIOR APPLICATION NUMBER: JP 2000-585
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-588
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-2299
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-26862
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: JP 2000-58367
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/JP00/09359
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 150
; SEQ ID NO 9
; LENGTH: 262
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-169-395-9

Query Match 46.2%; Score 1135; DB 15; Length 262;
Best Local Similarity 91.4%; Pred. No. 1.5e-89;
Matches 223; Conservative 2; Mismatches 7; Indels 12; Gaps 1;

QY 1 MTPEDPEETQPLIGPGGSAAPRRVFLAAFAAALGPLSGFALGYSSPAISLQRAAPP 60
DB 1 MTPEDPEETQPLIGPGGSAAPRRVFLAAFAAALGPLSGFALGYSSPAISLQRAAPP 60

QY 61 APRLDDAAASWFGAVVTGAAAGVVLGGWLVDRAGKLSLLCSVPFVAGFVITAAQDV 120
DB 61 APRLDDAAASWFGAVVTGAAAGVVLGGWLVDRAGKLSLLCSVPFVAGFVITAAQDV 120

QY 121 WMLLGGRLLTGLACGVASLVAPVYISEIAYPAVRGLLGSVQLMVVVGILLAYLAGWLE 180
DB 121 WMLLGGRLLTGLACGVASLVAPVYISEIAYPAVRGLLGSVQLMVVVGILLAYLAGWLE 180

QY 181 WRWLAIVGCVPSLSMLLMCFMPETPRFLLTQHRQEAALR 228
DB 181 WRWLAIVGCVPSLSMLLMCFMPETPRFLLTQHRQEAALR 240

QY 229 EOGW 232
DB 241 DPGW 244

RESULT 5
US-10-157-031-359
; Sequence 359, Application US/10157031
; Publication No. US20030108890A1
; GENERAL INFORMATION:
; APPLICANT: Baranova, A. V.
; APPLICANT: Yankovsky, N. K.
; APPLICANT: Kozlov, A. P.
; APPLICANT: Lobashev, A. V.
; APPLICANT: Krukovskaya, L. L.
; TITLE OF INVENTION: In silico screening for phenotype-associated expressed sequences
; FILE REFERENCE: 2760-103
; CURRENT APPLICATION NUMBER: US/10/157,031
; CURRENT FILING DATE: 2002-05-30
; NUMBER OF SEQ ID NOS: 415
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 359
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-157-031-359

Query Match 38.6%; Score 948; DB 14; Length 507;
Best Local Similarity 42.7%; Pred. No. 4.6e-73;
Matches 216; Conservative 80; Mismatches 166; Indels 44; Gaps 11;

QY 10 QPLIG-----PPGSAAPRGR-----RVFLAAPAAALGPLSGFALGYSSPA 50
DB 3 EPLLAGEGPDYDTFPEKPPSPGDRARVGTQNKRVFLATFAAVLGNFSGFALVYTSVP 62

QY 51 IPSLQRAAPPARLDDAAASWFGAVVTGAAAGVVLGGWLVDRAGKLSLLCSVPFVAG 110
DB 63 IPALERSLDDPDLHLTKSQASWFGSVFTLGAAGAGGSLAMTDLNLLGKLSIMFSAVSAAG 122

QY 111 FAVITAAQDVWMLLGGRLLTGLACGVASLVAPVYISEIAYPAVRGLLGSVQLMVVVGIL 170
DB 123 YALMAGAGHLMWLLLGRTLTGPAAGLTAACTIPVYVSEIAPPVGRGALGATPQLMAVFGSL 182

QY 171 LAYLAGWLEWRWLAIVGCVPSLSMLLMCFMPETPRFLLTQHRQEAALRFLWGSEQ 230
DB 183 SLVALGGLLPWRWLAIVAGEAPVLIIMLLSFMNPSRFLLSRGRDEALRALAWLRGTDV 242

QY 231 G--WEDPPIG---AEQSFHL--ALLARQGIYKPFIIIGVSLMAFQQLSGVNAVVFYAEITF 283
DB 243 DVHWEFEQIQDNVRRQSSRVSWAEARAPHVCHPITVALLMRLLQQTGTITPILVLIQSIF 302

QY 284 BEAKE-----KDSLSASVVVGVIVQLFTVAALIMDRAGRELLLVLSGVVNVFSTSAFGAY 339
DB 303 DSTAVLLLPKDD---AAIVGAVRLLSVLIAALTMDLAGRKVLLFVSAAMFAANLTGLY 359

QY 340 FKLTOGGPGNSSHVATSAFVS-----AQPVDSVGLAVGSMCLFIAGFVAGVGP 392
DB 360 IHF---GPRPLSPNSTAGLESSESWGDLAQPLAAPAGYLTLPVLLATMLTLMFIMGYAVGWGPI 416

QY 393 PHLLMSEIFPLHVKGATGICVLNTWMLMAFLVTKFSSLMELVLPYGAFWLASAFCLFSV 452
DB 417 TWLLMSEVLPLRARGVAGLCLVLAFLTAFLTKSFLPVVSTFGVLQVPPFFFAAICLVSL 476

QY 453 LFTFCVPETKGTLLEQITAFHE-GR 477
DB 477 VFTGCCVPETKGRSLEQIESFPTGR 502

RESULT 6

US-10-755-889-110
; Sequence 110, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10755,889
; PRIOR FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: Patent in version 3.2
; SEQ ID NO 110
; LENGTH: 507
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-110

Query Match 38.6%; Score 948; DB 16; Length 507;
Best Local Similarity 42.7%; Pred. No. 4.6e-73;
Matches 216; Conservative 80; Mismatches 166; Indels 44; Gaps 11;

QY 10 QPLG-----PPGSAFGR-----RVFLAAFAAALGPLSFGALGYSPA 50
Db 3 EPLGAGEDYDFPEKPPSPGDRARVGTLOKRVFLATFAVLGNFSGFVALVTSV 62

QY 51 IPSLQAAAPAPRLDDAAASWFGAVVTLGAAAGVGLGWLVDRAGRKLSLLCSVPFVAG 110
Db 63 IPALERSLPDLHLTKSQASWFSVFTLGAAGLSAMILNDLLGRKLSIMSAVPSAAG 122

QY 111 FAVITAAQDVMWLLGRLTLGLACGVASLVAPVYISEIAYPAVRGLLSCVQLMVGWGL 170
Db 123 YALMAGHGLMMLLGLRTLTGTGAGLTAACIPVYVSEIAPPVGRGALGATPOLMAVFGSL 182

QY 171 LAYLAGWLEWRWLAVLGCVPSPSLMLLMCFMPTPRFLLTQHRROEAMALRFLWGSEQ 230
Db 183 SLVALGGLLPWRWLAVAGAPVLMILLLSFNPSPRFLSRRDEEARLAWLRGTDV 242

QY 231 G--WEDPTG--AEQSHL--ALLRQPGIYKPFITIGVSLMAFQQLSGVNAVMYAETIF 283
Db 243 DVHWEPEIQDNVRRQSSRSVAEAPHCVRPITVALLMRLQLQTLGITPLVYLQSI 302

QY 284 BEAKF----KSSLASVVGVQVLEFVAALIMDRAGRLLLVLSGVVWFSTSAFGAY 339
Db 303 DSTAVLLPKDD---AAVGVARLLSVLIAATMDLAGRKVLLFVSAAIMFAANITLGLY 359

QY 340 FKLTOGGPNSHVAISAPVS-----AQPVDSVG-LAWLAVAGSMCLFIAGFVAGWGP 392
Db 360 IHF---GPRPLSPNSTAGLESSEWGDLAQPLAAGYTLVPLLATMFLFMINGYAGWGP 416

QY 393 PWMLMSEIFPLHVKGATGICVLTWMLMAFLVTKFSSLMELVRPYGAFWLASAPCIFSV 452
Db 417 TWLMLSEVLPLRARGVAGSLCVLASWLTAFVLTGKFLPVVSTFGVQVFFFAALCLVSL 476

QY 453 LFTLCVPTKTKTLEQITAFHE-GR 477
Db 477 VFTGCCVPTKTKSLEQIESFFRMGR 502

RESULT 7
US-10-051-909-38
; Sequence 38, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott

; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 38
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Oryza sativa
US-10-051-909-38

Query Match 29.4%; Score 723; DB 13; Length 501;
Best Local Similarity 35.6%; Pred. No. 1.2e-53;
Matches 173; Conservative 89; Mismatches 172; Indels 52; Gaps 8;

QY 17 GGSAPRGRRVFLAAFA-----AALGPLSFGALGYSSPAIPSLQRAAPPAP 62
Db 39 GGGGNGSRLLGSSAYSRLDSSVAVLCTLIIVALGPIQFGTCGSSPT----QDAIISDL 94

QY 63 RLDDAAASWFGAVVTLGAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFVITAAQDVMW 122
Db 95 GLTLESEFLSGLSNVGAMVGAISGQIAEYIGRKSLMIAAIPNIICWLAIISFAKDSF 154

QY 123 LLGRLTLGLACGVASLVAPVYISEIAYPAVRGLLSCVQLMVGWGLLAVLWLEWR 182
Db 155 LFMGRLEEGFGVGVISYVVPVYIABIAPQTMRGALGSVNLQSVTIGILLAYLGLMFPVR 214

QY 183 WLAVLGCVPSPSLMLLMCFMPTPRFLLTQHRROEAMALRFLWGSEQWEDPPIGAS-- 240
Db 215 ILSVLGILPCSLILPLGFFIPESPRWLAKMGKMGDEFSSQLVLRGFE-----TDIAEVN 269

QY 241 -----QSFHLALLRQPGIYKPFITIGVSLMAFQQLSGVNAVMYAETIFEEAKF 288
Db 270 EIKRSVQSSRRRTTIRFADIKKRYSVPLMVGIGLLVLQQLSGVNLIFYAASIFKAAGL 329

QY 289 KSSLASVVGVQVLEFVAALIMDRAGRLLLVLSGVVWFSTSAFGAYFKLTQGGPG 348
Db 330 TNSNLATPGLGVQVQVATGVTTWLTDRAGRRLLIISTGTITLVVVSVSF-FVKDNIT 388

QY 349 NSSHVAISAPVSAQPVDSVGLAWLAVAGSMCLFIAGFVAGWGPPIWLMSEIFPLHVKG 408
Db 389 NGSHL-----YSV-MSMLSLVGLVAFVIFSLGLGAIPWIMSEILPWNISL 435

QY 409 ATGICVLTWMLMAFLVTKFSSLMELVRPYGAFWLASAPCIFSVLFTLCVPTKTKTLE 468
Db 436 AGSVATLANMLTAWLITWT-ASLMLSNWNGTFAIYAACAGTILVFCVCLWVPTKTKRTLE 494

QY 469 QITAFH 474
Db 495 EIAFSF 500

RESULT 8
US-10-424-599-275264
; Sequence 275264, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 275264
; LENGTH: 484
; TYPE: PRT

```
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_90584C.1.pap
US-10-424-599-275264

Query Match      29.1%; Score 716; DB 15; Length 484;
Best Local Similarity 35.8%; Pred. No. 4.5e-53;
Matches 164; Conservative 89; Mismatches 174; Indels 34; Gaps 7;

QY 26 VFLAFAAALGPLSGFALGYSPPALPSLQRAAPPAPRLDDAAASWFGAVVTILGAAAGGV 85
DB 26 VFLAFAAALGPLSGFALGYSPPALPSLQRAAPPAPRLDDAAASWFGAVVTILGAAAGGV 85
QY 45 VLFCVLIVAGLPFOGFTGCGSSPTQGAIVRDL----NLISSEFFSGLSNVGAMVGAI 100
DB 45 VLFCVLIVAGLPFOGFTGCGSSPTQGAIVRDL----NLISSEFFSGLSNVGAMVGAI 100
QY 86 LGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDVNMLLGGRLLTGLACGVASLVAAPVYI 145
DB 86 LGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDVNMLLGGRLLTGLACGVASLVAAPVYI 145
QY 101 ASQIAEYIGRKSLMIAAIPNIIGWLAIISFAKSSFLYMGRLLEGFGVGIISYVVPVYI 160
DB 101 ASQIAEYIGRKSLMIAAIPNIIGWLAIISFAKSSFLYMGRLLEGFGVGIISYVVPVYI 160
QY 146 SETAYPAVRLGSGCVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 205
DB 146 SETAYPAVRLGSGCVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 205
QY 161 AEIAPQNLRLGSGSVNQLSVTIGIMLAYLLGLFVNVRLAILGILPCTVLIPGLFFIPES 220
DB 161 AEIAPQNLRLGSGSVNQLSVTIGIMLAYLLGLFVNVRLAILGILPCTVLIPGLFFIPES 220
QY 206 PRPLTQHRROEAMALRFLMGSEQGWEDPPI-----GAEQSFHLALLRQPGI 253
DB 206 PRPLTQHRROEAMALRFLMGSEQGWEDPPI-----GAEQSFHLALLRQPGI 253
QY 221 PRWLAKGMIDEPETSLQVLRGPD---TDSIEVHEIKRSVASTGKRAIRFADLKRKY 277
DB 221 PRWLAKGMIDEPETSLQVLRGPD---TDSIEVHEIKRSVASTGKRAIRFADLKRKY 277
QY 254 YPFIIGVSLMAFQQLSGVNAVMEFYAETIPFEAKFKDSSLASVYVGVIOVLFTVAALIM 313
DB 254 YPFIIGVSLMAFQQLSGVNAVMEFYAETIPFEAKFKDSSLASVYVGVIOVLFTVAALIM 313
QY 278 WFLPMVGILLVQLQSLGSGILFYSTTIFANAGISSEAAATVGLGAVQVIATGISTVLV 337
DB 278 WFLPMVGILLVQLQSLGSGILFYSTTIFANAGISSEAAATVGLGAVQVIATGISTVLV 337
QY 314 DRAGRLLLVLSGVVMVFSTSAFGAYFKLTQCGPNSHVAISAPVSAQPDVASVGLAW 373
DB 314 DRAGRLLLVLSGVVMVFSTSAFGAYFKLTQCGPNSHVAISAPVSAQPDVASVGLAW 373
QY 338 DKSGRLLLISSVNTVSLIIVSIAFYL-EGVVSDESHL-----PSILGIIVSIVGLVAM 391
DB 338 DKSGRLLLISSVNTVSLIIVSIAFYL-EGVVSDESHL-----PSILGIIVSIVGLVAM 391
QY 374 AVGSMCLFTIAGPAVGNGPIPLMLMSEIPPLHVKGVATGCVLTNMLMAFLVTKERSSLM 433
DB 374 AVGSMCLFTIAGPAVGNGPIPLMLMSEIPPLHVKGVATGCVLTNMLMAFLVTKERSSLM 433
QY 392 VI-----GFLGGLGPIPLMLMSEIPPLHVKGVATGCVLTNMLMSEIPPLHVKGVATGCVLTNMLM 443
DB 392 VI-----GFLGGLGPIPLMLMSEIPPLHVKGVATGCVLTNMLMSEIPPLHVKGVATGCVLTNMLM 443
QY 434 VLRPYGAFPLASAFICISVLFVLCVPTKGTLEQITAHF 474
DB 434 VLRPYGAFPLASAFICISVLFVLCVPTKGTLEQITAHF 474
QY 444 -WSSGGTFTIYTVAAFTIAFTIAFMWVPTKGTLEQITAHF 483
DB 444 -WSSGGTFTIYTVAAFTIAFTIAFMWVPTKGTLEQITAHF 483

RESULT 9
US-10-437-963-163544
; Sequence 163544, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 163544
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62530C.1.pap
US-10-437-963-163544

Query Match      28.6%; Score 702; DB 16; Length 501;
Best Local Similarity 35.9%; Pred. No. 7.5e-52;
Matches 166; Conservative 80; Mismatches 178; Indels 38; Gaps 6;

QY 26 VFLAFAAALGPLSGFALGYSPPALPSLQRAAPPAPRLDDAAASWFGAVVTILGAAAGGV 86
DB 26 VFLAFAAALGPLSGFALGYSPPALPSLQRAAPPAPRLDDAAASWFGAVVTILGAAAGGV 86
QY 85 FLCTLIIVAGLPFOGFTGCGSSPTQDAIIRDL----LTLSEFVSFSGLSNVGAMVGATA 140
DB 85 FLCTLIIVAGLPFOGFTGCGSSPTQDAIIRDL----LTLSEFVSFSGLSNVGAMVGATA 140
QY 87 GGMWVDRAGRKLSLLCSVPFVAGFAVITAAQDVNMLLGGRLLTGLACGVASLVAAPVYIS 146
DB 87 GGMWVDRAGRKLSLLCSVPFVAGFAVITAAQDVNMLLGGRLLTGLACGVASLVAAPVYIS 146
QY 141 SQMAEYIGRKSLMIAAIPNIIGWLAIISFAKSSFLYMGRLLEGFGVGIISYVVPVYIA 200
DB 141 SQMAEYIGRKSLMIAAIPNIIGWLAIISFAKSSFLYMGRLLEGFGVGIISYVVPVYIA 200
QY 147 ETAYPAVRLGSGCVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 206
DB 147 ETAYPAVRLGSGCVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 206
QY 201 EISPNMRGALGSVNQLSVTVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 260
DB 201 EISPNMRGALGSVNQLSVTVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 260
```

```
QY 27 FLAFAAALGPLSGFALGYSPPALPSLQRAAPPAPRLDDAAASWFGAVVTILGAAAGGV 86
DB 27 FLAFAAALGPLSGFALGYSPPALPSLQRAAPPAPRLDDAAASWFGAVVTILGAAAGGV 86
QY 63 FLCTLIIVAGLPFOGFTGCGSSPTQDAIIRDL----LTLSEFVSFSGLSNVGAMVGATA 118
DB 63 FLCTLIIVAGLPFOGFTGCGSSPTQDAIIRDL----LTLSEFVSFSGLSNVGAMVGATA 118
QY 87 GGMWVDRAGRKLSLLCSVPFVAGFAVITAAQDVNMLLGGRLLTGLACGVASLVAAPVYIS 146
DB 87 GGMWVDRAGRKLSLLCSVPFVAGFAVITAAQDVNMLLGGRLLTGLACGVASLVAAPVYIS 146
QY 119 SQMAEYIGRKSLMIAAIPNIIGWLAIISFAKSSFLYMGRLLEGFGVGIISYVVPVYIA 178
DB 119 SQMAEYIGRKSLMIAAIPNIIGWLAIISFAKSSFLYMGRLLEGFGVGIISYVVPVYIA 178
QY 147 ETAYPAVRLGSGCVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 206
DB 147 ETAYPAVRLGSGCVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 206
QY 179 EISPNMRGALGSVNQLSVTVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 238
DB 179 EISPNMRGALGSVNQLSVTVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 238
QY 207 RELTQHRROEAMALRFLMGSEQGWEDPPIGAE-----OSFHLALLRQPG 252
DB 207 RELTQHRROEAMALRFLMGSEQGWEDPPIGAE-----OSFHLALLRQPG 252
QY 239 RWLAKGMIDEPETSLQVLRGPD---TDSIAEVNDIKRAVASANKRTTIRFQELNOKK 293
DB 239 RWLAKGMIDEPETSLQVLRGPD---TDSIAEVNDIKRAVASANKRTTIRFQELNOKK 293
QY 253 IYKPFILGVSMAFQQLSGVNAVMEFYAETIPFEAKFKDSSLASVYVGVIOVLFTVAALI 312
DB 253 IYKPFILGVSMAFQQLSGVNAVMEFYAETIPFEAKFKDSSLASVYVGVIOVLFTVAALI 312
QY 294 YPFIIGVSLMAFQQLSGVNAVMEFYAETIPFEAKFKDSSLASVYVGVIOVLFTVAALI 353
DB 294 YPFIIGVSLMAFQQLSGVNAVMEFYAETIPFEAKFKDSSLASVYVGVIOVLFTVAALI 353
QY 313 MDRAGRLLLVLSGVVMVFSTSAFGAYFKLTQCGPNSHVAISAPVSAQPDVASVGLAW 372
DB 313 MDRAGRLLLVLSGVVMVFSTSAFGAYFKLTQCGPNSHVAISAPVSAQPDVASVGLAW 372
QY 354 LDRAGRRILLIISAGMTLSLLAVAVVFL-KDSISQDSHMYT-----LSM 399
DB 354 LDRAGRRILLIISAGMTLSLLAVAVVFL-KDSISQDSHMYT-----LSM 399
QY 373 LAVGSMCLFIAGPAVGNGPIPLMLMSEIPPLHVKGVATGCVLTNMLMAFLVTKERSSLM 432
DB 373 LAVGSMCLFIAGPAVGNGPIPLMLMSEIPPLHVKGVATGCVLTNMLMAFLVTKERSSLM 432
QY 400 ISLVALVAFVIAFSGMGAIPMIIMSEILPVSIGSLAGSFATLANWLTSGITMT-ANLM 458
DB 400 ISLVALVAFVIAFSGMGAIPMIIMSEILPVSIGSLAGSFATLANWLTSGITMT-ANLM 458
QY 433 EVLRPYGAFPLASAFICISVLFVLCVPTKGTLEQITAHF 474
DB 433 EVLRPYGAFPLASAFICISVLFVLCVPTKGTLEQITAHF 474
QY 459 LWSAGGTFTVSVVMSVSAFTLVFVILWVPETKGTLEQITAHF 500
DB 459 LWSAGGTFTVSVVMSVSAFTLVFVILWVPETKGTLEQITAHF 500

RESULT 10
US-10-425-114-63431
; Sequence 63431, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 63431
; LENGTH: 523
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4371-017-F11_FLI.pap
US-10-425-114-63431

Query Match      28.6%; Score 702; DB 15; Length 523;
Best Local Similarity 35.9%; Pred. No. 7.9e-52;
Matches 166; Conservative 80; Mismatches 178; Indels 38; Gaps 6;

QY 27 FLAFAAALGPLSGFALGYSPPALPSLQRAAPPAPRLDDAAASWFGAVVTILGAAAGGV 86
DB 27 FLAFAAALGPLSGFALGYSPPALPSLQRAAPPAPRLDDAAASWFGAVVTILGAAAGGV 86
QY 85 FLCTLIIVAGLPFOGFTGCGSSPTQDAIIRDL----LTLSEFVSFSGLSNVGAMVGATA 140
DB 85 FLCTLIIVAGLPFOGFTGCGSSPTQDAIIRDL----LTLSEFVSFSGLSNVGAMVGATA 140
QY 87 GGMWVDRAGRKLSLLCSVPFVAGFAVITAAQDVNMLLGGRLLTGLACGVASLVAAPVYIS 146
DB 87 GGMWVDRAGRKLSLLCSVPFVAGFAVITAAQDVNMLLGGRLLTGLACGVASLVAAPVYIS 146
QY 141 SQMAEYIGRKSLMIAAIPNIIGWLAIISFAKSSFLYMGRLLEGFGVGIISYVVPVYIA 200
DB 141 SQMAEYIGRKSLMIAAIPNIIGWLAIISFAKSSFLYMGRLLEGFGVGIISYVVPVYIA 200
QY 147 ETAYPAVRLGSGCVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 206
DB 147 ETAYPAVRLGSGCVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 206
QY 201 EISPNMRGALGSVNQLSVTVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 260
DB 201 EISPNMRGALGSVNQLSVTVGILLAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPET 260
```

QY 207 RFLLTQHRROBAMALRFLMGSEQWEDPPIGAE-----QSFHALLRQPG 252
Db 261 RWLAKNMDDFETSLQVLGRGF-----TDSAEVNDIKRAVASANKRTIRFQELNOKK 315
QY 253 IYKPFIIIGVSLMAFQQLSGVNAVNFYAEITFEERAKFKDSSLASVVGVIQVLTAVAALI 312
Db 316 YRTEILILGILLVLQQLSGINGILFYAGSIFKAAGLTNSDLATCALGAIQVLTATVWL 375
QY 313 MDRAGRILLVLGVVWVFTSAPGAYFKLTQGGPGNSHVAISAPVSAQPDASVGLAW 372
Db 376 LDRAGRILLIISAGMTLSLLAVAVVFFL-KDSISQDSHMYT-----LSM 421
QY 373 LAVGSMCLFIAGFVAGWGPPIWLLMSIFPLHVKGATGICVLTNWLMAFLVTKFESSLM 432
Db 422 ISLVALVAVFVAFSGNGALPWIMSEILFVSKSLAGSPATLANWLTSGITMT-ANLM 480
QY 433 EVLRPYGAFWLASAFICFSLVFLFCVPEPKGKTLQITAHF 474
Db 481 LWSAGTGFVSYVMVSAFTLVFLVWVPEPKGKTLQITAHF 522

RESULT 11
US-10-425-114-59933
; Sequence 59933, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingsong
; APPLICANT: Zhou, Yina
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 59933
; LENGTH: 509
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3689-227-G11_FLI pep
US-10-425-114-59933

Query Match 28.6%; Score 701.5; DB 15; Length 509;
Best Local Similarity 34.2%; Pred. No. 8.5e-52;
Matches 173; Conservative 86; Mismatches 182; Indels 65; Gaps 8;
QY 10 QPLLQ-----PPGSAAPRGR-----RVFLAFA-----AALGPLSRGF 42
Db 27 KPLINTGWNRPFPAGVWGVSQSSLMERLGSFSLRDVAISATLTLVAGLPQIFGF 86
QY 43 ALGYSSPAIPSIQRAAPPAPRLDDAAASWFGAVVTLGAAGVGLGWLVDRAKRLSLLL 102
Db 87 TCGYSSTP-----QDAIADLGLSEFSLFGLSNVGMVGAISQGLAEYIGKGLMI 142
QY 103 CSVPFVAGFAVITAQDVWMLLGGRLITGLACGVASLVAPVVISIAPVAVRGLGSCVQ 162
Db 143 AAIPIWIGLWLAISFKDGSFLFMGRLLGFGVGVISYTPVVIABIPQDORGALGSVQ 202
QY 163 LMVVVGIILAYLAGVLEWRWLVLCVPPSPSLMLLMCFMPTPRFLLTQHRROBAMAAL 222
Db 203 LSVTIGILLAYLFGFVFWRIILAVLILPCSLILIFGLFVFPSPWLAKMGWEDFEVSL 262
QY 223 RFLMGSEQWEDPPICAE-----QSFHALLRQPGIKYKPFIIIGVSLMAFQ 268
Db 263 QVLRGFQ-----TDIAEYNEIKRSLASSRRRTIRFADIKKRYSVPLVIGILLVQ 317
QY 269 LSGVNAVNFYAEITFEERAKFKDSSLASVVGVIQVLTAVAALMDRAGRILLVLGVV 328

Db 318 LSGVNGILFYAASIFKAAGITNSNLATFGLGAVQVATGVTTWLTDKAGRRLLLIISTTG 377
QY 329 MVFSTSAFGAYFKLTQGGPGNSHVAISAPVSAQPDASVGLAWLAVGSMCLFIAGFVAG 388
Db 378 MWITLVIVSVSFVKNDINAAAGSHLYSV-----MSMLSLAGLAVAFVAFSLG 423
QY 389 WGPPIWLLMSEIFPLHVKGATGICVLTNWLMAFLVTKFEFSSIMEVLRPYGAFWLASAF 448
Db 424 LGAIPIWIMSEILPWNKISLAGSVATLANWLTAWAITMT-ASLMLNWSSGGTFAIYAVVS 482
QY 449 IFSVLFTFLFCVPEPKGKTLQITAHF 474
Db 483 TWALIFVCLWVETKORTLEIAFSF 508

RESULT 12
US-10-051-909-36
; Sequence 36, Application US/10051909
; Publication No. US20020199217A1
; GENERAL INFORMATION:
; APPLICANT: Allen, Steve
; APPLICANT: Helentjaris, Tim
; APPLICANT: Hitz, Bill
; APPLICANT: Kinney, Tony
; APPLICANT: Tingey, Scott
; TITLE OF INVENTION: Plant Sugar Transport Proteins
; FILE REFERENCE: BB1163 US CIP
; CURRENT APPLICATION NUMBER: US/10/051,909
; CURRENT FILING DATE: 2002-01-17
; PRIOR APPLICATION NUMBER: 60/083,044
; PRIOR FILING DATE: April 24, 1998
; NUMBER OF SEQ ID NOS: 38
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 36
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: UNSURE
; LOCATION: (528)
; OTHER INFORMATION: Xaa = ANY AMINO ACID
US-10-051-909-36

Query Match 26.8%; Score 658.5; DB 13; Length 553;
Best Local Similarity 34.0%; Pred. No. 4.8e-48;
Matches 159; Conservative 77; Mismatches 189; Indels 43; Gaps 6;
QY 16 PGGSAPRGR-----VFLAFAAALGPLSGFALGYSPALPSLQRAAPPAPRLDDAAAS 70
Db 84 PGTSSMAVLRSHSVSAFLCTLIVAGLPQIFGFTSGFSPTQDAMVRDL-----NLSISEFS 139
QY 71 WFGAVVTLGAAGVGLGWLVDRAKRLSLLLCSVPFVAGFAVITAQDVWMLLGGRLIT 130
Db 140 AFGSLNVGVMGVAIGASQMAEYIGKGLMIATPNIIGWLAIISFAKDAFLYMGRLLE 199
QY 131 GLACGVASLVAPVVISIAPVAVRGLGSCVQVMVVGILLAYLAGVLEWRWLVLCV 190
Db 200 GFGVGLISYTPVVIABIPQDORGALGSVQNSVTFGIFLAYLLGMPFIMRLLAVAL 259
QY 191 PPSMLMLLMCFMPTPRFLLTQHRROBAMAALRFLMGSEQWEDPPICAE----- 240
Db 260 PCTMLIFGLFVFPSPWLAKMNLTEDCETSLQVLRGF-----TDITTEVNDIKRAVAS 314
QY 241 -----QSFHALLRQPGIKYKPFIIIGVSLMAFQQLSGVNAVNFYAEITFEERAKFKDSSLASV 296
Db 315 SSKRTTISFQELNOKKRYTPTLLIGILLVLQNLSGINGVLFIASSIFKAAGVITNSDLATC 374
QY 297 VVGVIQVLTAVAALIMDRAGRILLVLGVVWVFTSAPGAYFKLTQGGPGNSHVAIS 356
Db 375 SLGAIQVLTATVWLTDRAGRILLIISTSGMTCLLAVSVVFFLKDNLISQDSNSYI- 433
QY 357 APVSAQPDASVGLAWLAVGSMCLFIAGFVAGWGPPIWLLMSEIFPLHVKGATGICVLT 416

QY 180 ---EWRWLAVLG--CVPPSLMILLMCMFETPRFLLTQHRROBAMAALRFLMGSEQ-GWE 233
DB 179 ALNSGWRIPGLQLVPAALLLIGLLFLPESPRMLVEKGLKEAREVLAKLGVEDVDQE 238
QY 234 DPPIGAEQSFHLA-----LLR---QPGIYKPFIIIGVSLMAFOQLSGVNAVMFYA 279
DB 239 IQEIKAELEATVSEKAGKASWGELFRGTRPKVRQRLLMGVMQLQAFQQLTGAINAIFYYS 298
QY 280 ETIFEEAKFKD---SSLASVVVGVIOVLETAVAL-IMDRAGRRLLLVLSGVVMPFSTA 335
DB 299 PTIFKSVGVSDSVASLLVTIIVGVNVNFTFVALIFLVDVRRRRLLLGAAAGMAICFLI 358
QY 336 FGAYFKLTQGGPGNSHVAISAPVSAQPDVASVGLAWL-AVGSMCLFIAGFAVGMGPWP 394
DB 359 LGA-----SIGVALLLNKPKDPSKKAAGIVAIIVFILLFIAFFALGWPWP 405
QY 395 LLMSEIFPLHVKGVATGICVLTNWLMAFLVTKFSSLMELVRPY--GAFWLA---SAFCI 449
DB 406 VILSELFPKVRKSKALATAANWLANFII-----GFLFPYITGAIGLALGGVFLV 457
QY 450 FS-----VLFILFCVPETKGTLEQITAHF 474
DB 458 FAGLLVLFLFVFFVFPETKGTLEIEELF 488

RESULT 15

US-10-094-059-4
; Sequence 4, Application US/10094059
; Publication No. US20020127650A1
; GENERAL INFORMATION:
; APPLICANT: Millennium Pharmaceuticals Inc.
; TITLE OF INVENTION: 32468, A Human Sugar Transporter Family Member and
; TITLE OF INVENTION: Uses Therefor
; FILE REFERENCE: MPI01-040PIRM
; CURRENT APPLICATION NUMBER: US/10/094,059
; PRIOR FILING DATE: 2002-03-08
; PRIOR APPLICATION NUMBER: 60/275,053
; PRIOR FILING DATE: 2001-03-12
; NUMBER OF SEQ ID NOS: 8
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus
US-10-094-059-4

Query Match 25.0%; Score 615; DB 13; Length 488;
Best Local Similarity 35.0%; Pred. No. 2,3e-44;
Matches 179; Conservative 74; Mismatches 168; Indels 90; Gaps 18;
QY 30 AFAAALGPLSGFALGYSSPAIP-----SLQRAAPPAPRLDDAASWFGAVVT 77
DB 2 ALVAALGG--GFLFGYDGVIGGFLALIDFLFRGLLTSSGALAEVGYSTVLGLVVS 58
QY 78 ---LGAAGVGLGVLDVDRAGKLSLLCSVPFVAGFAVITAAQ-----DWW---MLGG 126
DB 59 IFFLGLIGSLFAGLGDGRFGRKSLIALVLFVIGALLSGAAGPYTTIGLWAFYLLIVG 118
QY 127 RLTLGLACGASLVAPVYIETAYPAVRGLGSCVQLMVMVCGILLAYLAGWVL----- 179
DB 119 RVLVGLGVGGASVLPWMIYSEIAPKALRGALGSLYQLAITIGILVAAIIGLNTNDS 178
QY 180 ---EWRWLAVLG--CVPPSLMILLMCMFETPRFLLTQHRROBAMAALRFLMGSEQ-GWE 233
DB 179 ALNSGWRIPGLQLVPAALLLIGLLFLPESPRMLVEKGLKEAREVLAKLGVEDVDQE 238
QY 234 DPPIGAEQSFHLA-----LLR---QPGIYKPFIIIGVSLMAFOQLSGVNAVMFYA 279
DB 239 IQEIKAELEATVSEKAGKASWGELFRGTRPKVRQRLLMGVMQLQAFQQLTGAINAIFYYS 298

QY 280 ETIFEEAKFKD---SSLASVVVGVIOVLETAVAL-IMDRAGRRLLLVLSGVVMPFSTA 335
DB 299 PTIFKSVGVSDSVASLLVTIIVGVNVNFTFVALIFLVDVRRRRLLLGAAAGMAICFLI 358
QY 336 FGAYFKLTQGGPGNSHVAISAPVSAQPDVASVGLAWL-AVGSMCLFIAGFAVGMGPWP 394
DB 359 LGA-----SIGVALLLNKPKDPSKKAAGIVAIIVFILLFIAFFALGWPWP 405
QY 395 LLMSEIFPLHVKGVATGICVLTNWLMAFLVTKFSSLMELVRPY--GAFWLA---SAFCI 449
DB 406 VILSELFPKVRKSKALATAANWLANFII-----GFLFPYITGAIGLALGGVFLV 457
QY 450 FS-----VLFILFCVPETKGTLEQITAHF 474
DB 458 FAGLLVLFLFVFFVFPETKGTLEIEELF 488

Search completed: February 24, 2005, 06:04:07
Job time : 134 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 05:44:20 ; Search time 44 Seconds
(without alignments)
1043.077 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPEPQPLGPPGSA.....CVPETKGTLEQITAHPEGR 477

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 483416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:.*
1: pir1:.*
2: pir2:.*
3: pir3:.*
4: pir4:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	734	29.9	487	2 E96782	hypothetical prote
2	722	29.4	490	2 T14545	probable sugar tra
3	607	24.7	461	2 D70073	metabolite transpo
4	607	24.7	471	2 AB0868	L-arabinose isomer
5	606.5	24.7	472	2 S47089	arabinose-proton s
6	603	24.5	472	2 B91091	L-arabinose isomer
7	603	24.5	472	2 E85936	L-arabinose isomer
8	603	24.5	472	2 B26430	L-arabinose isomer
9	591.5	24.1	464	2 F65079	galactose-proton s
10	587.5	23.9	464	2 C91106	galactose-proton s
11	587.5	23.9	464	2 F85951	galactose-proton s
12	584.5	23.8	496	2 T52132	probable sugar tra
13	584.5	23.8	575	2 T43400	myo-inositol trans
14	583.5	23.7	464	2 AC0877	galactose-proton s
15	582	23.7	521	2 G84864	probable membrane
16	574.5	23.4	493	2 S38981	glucose transport
17	570.5	23.2	496	2 A31986	glucose transport
18	569.5	23.2	493	2 A41751	glucose transport
19	567	23.1	580	2 D86426	hypothetical prote
20	560.5	22.8	457	2 E70070	metabolite transpo
21	556.5	22.6	560	2 T51485	sugar transporter-
22	554.5	22.6	464	2 F69587	L-arabinose transp
23	551	22.4	461	2 G85059	probable sugar tra
24	548.5	22.3	523	2 S25015	monosaccharide tra
25	546.5	22.2	526	2 T01853	probable hexose tr
26	546	22.2	491	2 B86096	xylose-proton symp
27	546	22.2	491	2 F91255	xylose-proton symp
28	546	22.2	491	2 A26430	xylose transport p
29	542.5	22.1	522	2 A31556	glucose transport

ALIGNMENTS

RESULT 1

E96782
hypothetical protein F22H5.6 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cross)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
C:Accession: E96782
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso ansen, N.F.; Hughes, B.; Huizar, L.
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: E96782
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-487 <STO>
A:Cross-references: UNIPROT:Q9FRL3; GB:AE005173; NID:g10092276; PIDN:AAG12689.1; GSPDB: C:Genetics:
A:Gene: F22H5.6
A:Map position: 1
C:Superfamily: glucose transport protein

Query Match 29.9%; Score 734; DB 2; Length 487;
Best Local Similarity 37.5%; Pred. No. 1.1e-45;
Matches 173; Conservative 77; Mismatches 177; Indels 34; Gaps 6;

QY 26 VFLAFAAALGSLGFGALGYSIPAIPSLQRAAPAPRLDDAAASWFGAVTLGAAGGV 85
DB 48 VLACVLIVAGLGIQFGTCGYSPT---QAATKDLGLTVEYSVFGSLNVGAVGAI 103
QY 86 LGGLVLDVDRAGKSLLLCSVPFVAGFVAVITAAQDVNMLGGRLTLTCLAGVASLAVPVYI 145
DB 104 ASGQTAIEYGRKGLMIAIPNIIGLCSFPAKDTFLYNGRLLEGFGVGIISYTPVVI 163
QY 146 SEIAYPAVRGLGSCVQLMVVVGILLAYLAGVLEWRVLAVLGCVPSPSLMLLMCFMPE 205
DB 164 AEIAQNMREGGLSGVNLQSVTIGIMLAYLLGLFVFWRLAVLGLPCTLLIPGLPIFES 223
QY 206 PRLFLTQHRQRQAMALRFLMGSEOGWEDPPGAEQ-----SFHLALLRQPGI 253
DB 224 PRLARQMGMTDEFSLQVLGRFE---TDITVEVNEIKESVASSTKRNTRVRFDLKRRY 280
QY 254 YKPEITIGVSLMAFQQLSGVNAVMEYAEITIFERAKFKDSSLASVWGVIOVLFTFAVALIM 313
DB 281 YPFLMWGVLVLQQLGGINGVLFTSYSTIFESAGVTSSNAATFGVGAIQVATAISTWL 340
QY 314 DRAGRRLLLVLSGVVMVFSTSAFGAYFKLTGGPGNGSHVAISAPVSAQFVDASVGLAWL 373

C;Accession: D70073
R;Kunst, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.
koetter, P.; Konigstein, G.; Krogh, S.; Kunano, M.; Kurita, K.; Lapidus, A.; Lardinois
A;Authors: Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, K.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yaumoro, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
C;Accession: D70073
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-461 <KUN>
A;Cross-references: UNIPROT:P46333; GB:Z99124; GB:AL009126; NID:g2636442; PIDN:CAB16017.
A;Experimental source: strain 168
C;Genetics:
A;Gene: yxcC
C;Superfamily: glucose transport protein

Query Match 29.4%; Score 722; DB 2; Length 490;
Best Local Similarity 36.1%; Pred. No. 8.1e-45;
Matches 167; Conservative 81; Mismatches 177; Indels 38; Gaps 5;

QY 26 VFLAFAAALGPLSGFALGYSSPAIPSLQRAAPPAPRLDDAAASWFGAVVTLGAAGGV 85
DB 50 VLACVLVALGPIQFOFTAGYSPT---QSATNELGSLVAYEYNSFGLSVNGAVGAI 105
QY 86 LGGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDVWMLLGGRLTLGLACVASLAPVYI 145
DB 106 ASQISEYIGRKSLMIAIPNIIGLWLAISFAKDSFLYMGRLGFGVGIISYTPVYI 165
QY 146 SETAYPAVRGLLSCVOLMVVVGILLAYLAGVLEWLVAVLGCVPSPSLMLLMCFMET 205
DB 166 SETAPQNLRGALGSVNQLSVTIGIMLSYMLGLFVPWRLAVILGILPCTILPLGLFIPES 225
QY 206 PRELLTQHRROEAMALRFLWSEQGWEDPPICAE-----QSPHLALLROP 251
DB 226 PRLAKWGMEEPEVSLQVLRG----DTDLSLEWNEIKRSVASSKETTIRFAELRQR 280
QY 252 GIYKPIIIGVSLMAFQOLSGVNAVMPYAEITIFEEKFKDSSLASVVVGVIQVLTFAVAL 311
DB 281 RYWLPLMIGNLLIQLQSLGVLFSYTSITKEAGVTSNSAATFGLGAVQVIATVITW 340
QY 312 INDRAGRLLLVLSGVVVFSTSAFGAYFKLTQGGPGNSHVAISAPVSAQPDASVGLA 371
DB 341 LVDKSGRLLLVSSSGMTLSLLVWMSFPLEKMSVDSDESTYSV-----FS 386
QY 372 WLAVGSMCLFIAGFVAGWGPPIWLLMSEIFPLHVKGVATGICVLTNWLMFLVTKFESSL 431
DB 387 ILSVGVVNAVVTFSIGIAPWIMSEILPINIKLAGSIATLANWFVAVITWT-ANI 445
QY 432 MEVLRYPGAFWLASAFICFISVLTFLFCVPETGKTLQITAHF 474
DB 446 MLSWNSGGTFSIYMWCAFTVAFWVWVWVETKGTLEELIOWSF 488

RESULT 3
D70073
metabolite transport protein homolog yxcC - *Bacillus subtilis*
C;Species: *Bacillus subtilis*
C;Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C;Accession: D70073
R;Kunst, P.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Bruch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Cho
A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A;Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galler
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holzapfel, S.; Hosono, S.; Hullo, M.F.
koetter, P.; Konigstein, G.; Krogh, S.; Kunano, M.; Kurita, K.; Lapidus, A.; Lardinois
Y. M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, M.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A;Authors: Schleich, S.; Schroeter, R.; Scoffone, P.; Sekiguchi, J.; Sekowska, A.; Seror
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, K.; Tosato, V.; Uchiyama
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yaumoro, K.; Yata, K.; Yoshida, K.
A;Authors: Yoshikawa, H.F.; Zumbstein, E.; Yoshikawa, H.; Danchin, A.
A;Title: The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*.
A;Reference number: A69580; MUID:98044033; PMID:9384377
C;Accession: D70073
A;Status: preliminary; nucleic acid sequence not shown; translation not shown
A;Molecule type: DNA
A;Residues: 1-461 <KUN>
A;Cross-references: UNIPROT:P46333; GB:Z99124; GB:AL009126; NID:g2636442; PIDN:CAB16017.
A;Experimental source: strain 168
C;Genetics:
A;Gene: yxcC
C;Superfamily: glucose transport protein

Query Match 24.7%; Score 607; DB 2; Length 461;
Best Local Similarity 31.1%; Pred. No. 1.6e-36;
Matches 146; Conservative 85; Mismatches 192; Indels 46; Gaps 7;

QY 25 RVFLAFAAALGPLSGFALGYSSPAIPSLQRAAPPAPRLDDAAASWFGAVVTLGAAGG 84
DB 6 RXYMIYFPGAUGLGYGYDTGIVSGALLFNNDIPLTLTEGLVVS---MLLGLAFGS 61
QY 85 VLGGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDVWMLLGGRLTLGLACVASLAPVYI 144
DB 62 ALSGTCSDRWGRKRVVFLSIIFIIGALACAFSQTIGMLIASRVILGLAVGGSTALVPVY 121
QY 145 ISETAYPAVRGLLSCVOLMVVVGILLAYLAGVWL----EWRWLAVLGCVPSPSLMLLMCM 200
DB 122 LSEMAPTKIRGTGLTMNLMIVTGLLAIYVNYLFTFPEARWVWVGLAAVAVLLLGIA 181
QY 201 FMPTETPRLTQHRROEAMALRF-----LWSEQGWEDPPICAEOSFHALLRQ 250
DB 182 FMPSRPLWLVKRGSEERARINWTHDPKDIEMELAEKQGEAE-----KKEITLGVLKA 236
QY 251 PGYKPIIIGVSLMAFQOLSGVNAVMPYAEITIFEEKF--KDSLASVVVGVIQVLTFAV 308
DB 237 KWIPLMLLIGVGLAIFQQAVGINTVIYVYPTFTKAGLGTSSALGTMGILNVIMCIT 296
QY 309 AALMDRAGRLLLVLSGVVVFSTSAFGAYFKLTQGGPGNSHVAISAPVSAQPDASV 368
DB 297 AMILIDRVGRKLLIWSVGITLSLAALSGVL-LTLG-----LSA 335
QY 369 GLAVLAVGSMCLFIAGFVAGWGPPIWLLMSEIFPLHVKGVATGICVLTNWLMFLVTKFE 428
DB 336 STAMTVVFLGVYIVFQATGVPVWVLMPELFSKARGATGTTLVLSAANLVSLVF 395
QY 429 SSLMEVLRYPGAFWLASAFICFISVLTFLFCVPETGKTLQITAHFEGR 477
DB 396 PLMLSAMGIAWVWVVFSLCLSSFFFAFYVMPVETKGSLEIEASLKKR 444

RESULT 4
AB0868
L-arabinose isomerase [imported] - *Salmonella enterica* subsp. *enterica* serovar Typhi (st
C;Species: *Salmonella enterica* subsp. *enterica* serovar Typhi
A;Note: this species has also been called *Salmonella typhi*
C;Date: 09-Nov-2001 #sequence_revision 09-Nov-2001 #text_change 18-Nov-2002
C;Accession: AB0868
R;Parkhill, J.; Dougan, G.; James, K.D.; Thomson, N.R.; Pickard, D.; Wain, J.; Churcher,
th, T.; Connerton, P.; Cronin, A.; Davis, P.; Davies, R.M.; Dowd, L.; White, N.; Farrar,

[illegible]

Query Match	24.7%;	Score	606.5;	DB 2;	Length	472;	
Best Local Similarity	33.7%;	Pred.	No. 1.8e-36;				
Matches	153;	Conservative	84;	Mismatches	182;	Indels	35; Gaps
8;							
QY	30	AFAAALGPLSFGFALGVSSPAISLQRAAPAPRLDDAASWFGAVVTTLGAAAGVLGSGW	89				
DB	26	SIANAAGLLFGDIDGVIAGALPFTDHFVLSRLQE-----WVSSMMGLGAAGALFNGW	81				
QY	90	LVDRAGRKLSLLCSVPFVAGFAVITAAQDVMMLLGGRLLTLGLACGVASLVAPVYISRIA	149				
DB	82	LSFRLGRKYSLVAGVILFVAGSVCSAFATSVEMLLVARIVLGVAGVIASTYAPLYLSEWA	141				
QY	150	YPAVRGLGSCVOLMVVVGILLAYLAGWLEW--RWLAVLG--CUPSMLMLLMCFMPEET	205				
DB	142	SENVGRKMTSMYQLMVTTLGIVMAFLSDTAFSYSGNWRAMLGLVALPAVVLIIILVIFLPS	201				
QY	206	PRELLTQHRQREAAALRFLMGSEOGWEDPPIGAEQSFHL-----ALLR-OPGLVKPFI	259				
DB	202	PRWLAEKGRHVEAEVLRLMDTSEKARDELNIRESLKLGQGWALFKVNRNVRRAVFL	261				
QY	260	GVSLMAFPQGLSGYNVAFYATIEFAK---KDSSLASVVVGVIQVLFTTAAALIMDRA	316				
DB	262	GMLLQAMQOFTGNNIIMYAPRIFKMAGFTTTEQOMVATLVVGLTFEAFATIAVFTVDKA	321				
QY	317	GRRLILVLSGVNVFSTSAFGVKLTQGGPGNSHVAISAPVSAQPDVDSVGLAWLAVG	376				
DB	322	GRKPALKIGFSVMAIGTLVLG--YCLMQPDNGT-----ASSGLSLWSVG	363				
QY	377	SMCLFTAGFVAGCGPIPWLLMSIFPLHVKGVATGICVLTNMLMAFLVTKFPSSIMEVLR	436				
DB	364	MTWMCITAGYAMSAPVVMILCSIQPKRDFGTCITTTNNWSNMIIIGATFLTLDDAIG	423				
QY	437	PYGAFWLASAFICFSVLTLFCVPETKGGTKLEQI	470				
DB	424	AACTFWLYTALNVAFICGVFWLIPETKQVLEHI	457				
RESULT	6						
B91091							
L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli (strain O157:H7, substrain RIMD 0356:1)							
C:Species: Escherichia coli							
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004							
C:Accession: B91091							
R:Hayashi, T.; Makino, K.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C. G. S. S.; Sasagawa, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hattori, M.; Shinagawa, H. DNA Res. 8, 11-22, 2001							
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genomic islands							
A:Reference number: A99629; MUID:21156231; PMID:11258796							
A:Accession: B91091							
A:Status: preliminary							
A:Molecule type: DNA							
A:Residues: 1-472 <HAY>							
A:Cross-references: UNIPROT:P09830; GB:BA000007; PIDN:BAB37121.1; PID:gl3363170; GSPDR:BA000007							
A:Experimental source: strain O157:H7, substrain RIMD 0509952							
C:Genetics:							
A:Gene: EC83698							
C:Superfamily: Glucose transport protein							
C:Keywords: intramolecular oxidoreductase; isomerase							
Query Match	24.5%;	Score	603;	DB 2;	Length	472;	
Best Local Similarity	33.3%;	Pred.	No. 3.1e-36;				
Matches	158;	Conservative	84;	Mismatches	197;	Indels	36; Gaps
9;							
QY	9	TQPLIGPPGSGRGRVFLAAFAAALGPLSGFALGVSSPAISLQRAAPAPRLDDAA	68				
DB	6	TSALTPRSRLDTRRNMMFVSV--AAAVAGLFGDIDGVIAGALPFTDHFVLSRLQE--	62				
QY	69	ASWFGAVVTTLGAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDVMMLLGGRLL	128				
DB	63	--WVSSMMGLGAAGALFNGWLSFRLGRKYSLMAGALLFVLGSLGSAFATSVEMLIAARV	120				
QY	129	LTGLACGVASLVAPVYISEIAYPAVRGLGSCVOLMVVVGILLAYLAGWLEW--RWLAV	186				

Db 121 VLGIAGIASYATAPLYLSEMAENVRGKMSIYQMLVTLGIVLAFSLDTAFSYSGNWRAM 180
QY 187 LG--CVPPSLMLLLMCFPEPTRFLTLTQHRROEAAALRFLWGSQGWEDPPIGAQSFH 244
Db 181 LGVLALPAVLLIILVFLFNSPRWLAEGRHTEAEVEVLRMLRDTSEKAREELNEIRESLK 240
QY 245 L-----ALLR-QPGIVKPIIGVSLMAFOOLSGVNAVVFYAEETIPEEAKF---KDSLSAS 295
Db 241 LKQGGWALFKINVRRAVFLGMLLQAMQOFTGMNIIYAPRIFKMAGFTTTEOQMIAT 300
QY 296 VVVGVIQVLTAAALIMDRAGRRLLVLSGVVMVFSTSAFGAYFKLTQGGPGNSSHVAI 355
Db 301 LVVGLTFMFATFIATFVTDKAGRKPKALKIGFSVMAIAGTLVLG--YCLMQFDNGT----- 352
QY 356 SAPVSAQPDVASGLAWLAVGSKCLFIAGFVAGWGPPIWMLMSEIFPLHVKGVATGICVL 415
Db 353 -----ASSGLSWLSVGMTMCMCIAGYAMSAAPVVMWILCSEIQPLKCRDFGTCSTT 402
QY 416 TNWLMAPLVTKBFSSLMELVRPVGAFWLASAFICFVSFLTFLCVPETKGTLEQI 470
Db 403 TNWVSNMIIATFTLLDSIGAAGTFWLTALNIAFVGITFWLIPETKKNVLEHI 457

RESULT 7
E85936
L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli (strain O157:H7, substrain EDL933)
C:Species: Escherichia coli
C>Date: 16-Feb-2001 #sequence_revision 16-Feb-2001 #text_change 09-Jul-2004
C:Accession: E85936
R:Perna, N.T.; Plunkett III, G.; Burland, V.; Mau, B.; Glasner, J.D.; Rose, D.J.; Mayhew
iller, L.; Grobeck, E.J.; Davis, N.W.; Lim, A.; Dimalanta, E.; Potamousis, K.; Apodaca,
Nature 409, 529-533, 2001
A:Title: Genome sequence of enterohemorrhagic Escherichia coli O157:H7.
A:Reference number: A85480; MUID:21074935; PMID:11206551
A:Accession: E85936
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-472 <STO>
A:Cross-references: UNIPROT:P09830; GB:AE005174; NID:G12517333; PIDN:AG57953.1; GSPDB:C
A:Experimental source: strain O157:H7, substrain EDL933
C:Genetics:
A:Gene: araE
C:Superfamily: glucose transport protein
C:Keywords: Intramolecular oxidoreductase; isomerase

Query Match 24.5%; Score 603; DB 2; Length 472;
Best Local Similarity 33.3%; Pred. No. 3.1e-36;
Matches 158; Conservative 84; Mismatches 197; Indels 36; Gaps 9;

QY 9 TQPLGPPGGSAPGRRRVFLAAFAAALGPLSGFALGYSSPAIFSLQRAAPPAPRLDDAA 68
Db 6 TESALTFRSLDRTRMNMFSV-AAAAGLLFGLDIGVIAGALPFTIDHFVLTSLRQE-- 62

QY 69 ASWFGAVVTLGAAAGVGLGWLVDRAGKLSLLCSVPFVAGFAVITAAQDVWMLLGGRL 128
Db 63 --VWVSSMMLGAAIGALFNGWLSFRLGRKYSLMAGAILFVLGSGSAFATSEVMLIAARV 120

QY 129 LTGLACGVASLAPVPISEIAYPAVRGLGSCVOLMVVVGILLAYLAGWLEW--RWLAV 186
Db 121 VLGIAGIASYATAPLYLSEMAENVRGKMSIYQMLVTLGIVLAFSLDTAFSYSGNWRAM 180

QY 187 LG--CVPPSLMLLLMCFPEPTRFLTLTQHRROEAAALRFLWGSQGWEDPPIGAQSFH 244
Db 181 LGVLALPAVLLIILVFLFNSPRWLAEGRHTEAEVEVLRMLRDTSEKAREELNEIRESLK 240

QY 245 L-----ALLR-QPGIVKPIIGVSLMAFOOLSGVNAVVFYAEETIPEEAKF---KDSLSAS 295
Db 241 LKQGGWALFKINVRRAVFLGMLLQAMQOFTGMNIIYAPRIFKMAGFTTTEOQMIAT 300

QY 296 VVVGVIQVLTAAALIMDRAGRRLLVLSGVVMVFSTSAFGAYFKLTQGGPGNSSHVAI 355
Db 301 LVVGLTFMFATFIATFVTDKAGRKPKALKIGFSVMAIAGTLVLG--YCLMQFDNGT----- 352

QY 356 SAPVSAQPDVASGLAWLAVGSKCLFIAGFVAGWGPPIWMLMSEIFPLHVKGVATGICVL 415
Db 353 -----ASSGLSWLSVGMTMCMCIAGYAMSAAPVVMWILCSEIQPLKCRDFGTCSTT 402

QY 416 TNWLMAPLVTKBFSSLMELVRPVGAFWLASAFICFVSFLTFLCVPETKGTLEQI 470
Db 403 TNWVSNMIIATFTLLDSIGAAGTFWLTALNIAFVGITFWLIPETKKNVLEHI 457

RESULT 8
B26430
L-arabinose isomerase (EC 5.3.1.4) - Escherichia coli (strain K-12)
C:Species: Escherichia coli
C>Date: 05-Oct-1988 #sequence_revision 05-Oct-1988 #text_change 09-Jul-2004
C:Accession: B26430; A28075; I40996; B65067
R:Maiden, M.C.J.; Davis, E.O.; Baldwin, S.A.; Moore, D.C.M.; Henderson, P.J.F.
Nature 325, 641-643, 1987
A:Title: Mammalian and bacterial sugar transport proteins are homologous.
A:Reference number: A93389; MUID:87115869; PMID:3543693
A:Accession: B26430
A:Status: nucleic acid sequence not shown; not compared with conceptual translation
A:Molecule type: DNA
A:Residues: 1-472 <MAI>
A:Cross-references: UNIPROT:P09830
R:Maiden, M.C.J.; Jones-Mortimer, M.C.; Henderson, P.J.F.
J. Biol. Chem. 263, 8003-8010, 1988
A:Title: The cloning, DNA sequence, and overexpression of the gene araE coding for arabi
A:Reference number: A28075; MUID:88228015; PMID:2836407
A:Accession: A28075
A:Molecule type: DNA
A:Residues: 1-472 <MA2>
A:Cross-references: GB:J03732; NID:G145320; PIDN:AAA23469.1; PID:G145321
R:Stoner, C.; Schleif, R.
J. Mol. Biol. 171, 369-381, 1983
A:Title: The araE low affinity L-arabinose transport promoter. Cloning, sequence, trans
A:Reference number: I40996; MUID:84114868; PMID:6319708
A:Accession: I40996
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-25, 'Y', 349, 'R', <RES>
A:Cross-references: EMBL:X00272; NID:G40940; PIDN:CAA25075.1; PID:G40941
R:Blattner, F.R.; Plunkett III, G.; Bloch, C.A.; Perna, N.T.; Burland, V.; Riley, M.; Co
.A.; Rose, D.J.; Mau, B.; Shao, Y.
Science 277, 1453-1462, 1997
A:Title: The complete genome sequence of Escherichia coli K-12.
A:Reference number: A64720; MUID:97426617; PMID:9278503
A:Accession: B65067
A:Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-472 <BLAT>
A:Cross-references: GB:AB000368; GB:U00096; NID:G2367165; PIDN:AAC75880.1; PID:G1789207;
A:Experimental source: strain K-12, substrain MG1655
C:Genetics:
A:Gene: araE
A:Map position: 61 min
C:Superfamily: glucose transport protein
C:Keywords: arabinose transport; intramolecular oxidoreductase; isomerase; membrane prot

Query Match 24.5%; Score 603; DB 2; Length 472;
Best Local Similarity 33.3%; Pred. No. 3.1e-36;
Matches 158; Conservative 84; Mismatches 197; Indels 36; Gaps 9;

QY 9 TQPLGPPGGSAPGRRRVFLAAFAAALGPLSGFALGYSSPAIFSLQRAAPPAPRLDDAA 68
Db 6 TESALTFRSLDRTRMNMFSV-AAAAGLLFGLDIGVIAGALPFTIDHFVLTSLRQE-- 62

QY 69 ASWFGAVVTLGAAAGVGLGWLVDRAGKLSLLCSVPFVAGFAVITAAQDVWMLLGGRL 128
Db 63 --VWVSSMMLGAAIGALFNGWLSFRLGRKYSLMAGAILFVLGSGSAFATSEVMLIAARV 120

QY 129 LTGLACGVASLAPVPISEIAYPAVRGLGSCVOLMVVVGILLAYLAGWLEW--RWLAV 186
Db 121 VLGIAGIASYATAPLYLSEMAENVRGKMSIYQMLVTLGIVLAFSLDTAFSYSGNWRAM 180

Nature 409, 529-533, 2001

A;Title: Genome sequence of enterohemorrhagic *Escherichia coli* O157:H7.
A;Reference number: A85480; MUID:21074935; PMID:11206551
A;Accession: F85951
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-464 <STO>
A;Cross-references: UNIPROT:O8XCW7; GB:A8005174; NID:g12517486; PIDN:AAG58074.1; GSPDB:G12517486; strain O157:H7, substrate EDL933
C;Genetics:
C;Gene: Galp
C;Superfamily: glucose transport protein

Query Match 23.8%; Score 587.5; DB 2; Length 464;
Best Local Similarity 32.2%; Pred. No. 4.1e-35;
Matches 154; Conservative 83; Mismatches 176; Indels 65; Gaps 11;

QY 27 FLAFAAALGFLSFGFALGYSPPAIPSLQRAAPPAPFLDDAASWFGAVNTLGAAGCVL 86
DB 16 FFCVFLAALAGLFLGDLGIVAGLPIADEF---QITSHQEWVSSMMFGAAGVAVG 71
QY 87 GGMVLDRAGRKLSLLCSVPFVAGFVITAAQDVMMLLGGRLLTGLACGVASLVAPVYIS 146
DB 72 SGWLSFKLGRKSLMTGAILFVAGLSFSAAPPNVEVLILSRVLLGLAVGASTAPLYLS 131
QY 147 EIAYPVAVRGLGSCVQLMVVVGILLAYLA---GWLEWRWLAVLGCVPSPSLMLLMCFM 202
DB 132 EIAPEKIRGSMISWYQLMITIGILGAYLSDTAFSTGAWRWMLGVIIIPAILLLIGVFFL 191
QY 203 PETPRFLLTOHRQEAVAL-----REFMGSEQWEDDPDGAQSFH 244
DB 192 PDSFPPFAAKRRFVDAERVLRLRDTSAEAKRELDEIRSLQKSGW----- 239
QY 245 LALLRQPGIYKPT-IGVSLMAFQOLSGVNAVVFYAEIIEEAKFKDSS---LASVVVGV 300
DB 240 -ALFKNSNFRAVFLGVLQVQCFQGMVMIYAPKIFELAGYTNITTEQWNGTVIVGL 298
QY 301 IOVLFTAAVALIMDRAGRRLLLVLSGVVWVFSTSAFGAYFKLQGGPGNSHVAISAPVS 360
DB 299 TNVLATFIALGLVDRWGRKPTLTGLFVW-----MGLGTMMHIGIHP-S 346
QY 361 AQPVDASVGLAWLAVGSMCLFIAGFVAGVGPDPWLLMSEIFPLHVKGVATGICVLINWLM 420
DB 347 AQ-----YFAIAMLMLFVIGFAMSGFLWLCSEIQPLKGRDFGTCSTATNWIA 397
QY 421 AFLVTKFEFSLMEVLRPYGAFWLASAFICISVLTFLFCVPETGKTLQIQTAAH-PFGR 477
DB 398 NMIVGATFLMLNTLGNANTFWAGLNVLFIILLVLPETKRVSLHEIRNLMKGR 455

RESULT 12
T52132
Probable sugar transporter protein ERD6 [imported] - Arabidopsis thaliana
C;Species: Arabidopsis thaliana (mouse-ear cress)
C;Date: 20-Oct-2000 #sequence_revision 20-Oct-2000 #text_change 09-Jul-2004
C;Accession: T52132
R;Kiyosue, T.; Abe, H.; Yamaguchi-Shinozaki, K.; Shinozaki, K.
Biochim. Biophys. Acta 1370, 187-191, 1998
A;Title: ERD6, a cDNA clone for an early dehydration-induced gene of Arabidopsis, encode
A;Reference number: Z25973; MUID:98213606; PMID:9545564
A;Accession: T52132
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-496 <KIT>
A;Cross-references: UNIPROT:O65799; EMBL:D89051; PIDN:BAA25989.1
C;Genetics:
A;Gene: ERD6
C;Superfamily: glucose transport protein
C;Keywords: transmembrane protein; transport protein

Query Match 23.8%; Score 584.5; DB 2; Length 496;
Best Local Similarity 33.5%; Pred. No. 7.2e-35;
Matches 160; Conservative 73; Mismatches 169; Indels 75; Gaps 11;

Db 434 NWISNLIVAQTFLTIAAAGTGMTFLILAGIAVLAVIFVIVFVPETQGLTFSEV 487

Search completed: February 24, 2005, 05:53:46
Job time : 46 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 05:43:22 ; Search time 180 Seconds
(without alignments)
1357.010 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPDPETQPLLPGGSA.....CVPETKGTLEITAHFGR 477

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Uniprot_03.*
1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2457	100.0	477	Q8WU29	Q8wuz9 homo sapien
2	2445	99.5	477	1 GTR8 HUMAN	Q9ny64 homo sapien
3	2230.5	90.8	478	2 Q6XU12	Q6xui2 bos taurus
4	2162.5	88.0	478	1 GTR8 RAT	Q9jjz1 rattus norv
5	2156	87.7	477	1 GTR8 MOUSE	Q9jjf3 mus musculu
6	1537.5	62.6	334	1 GTR8 BOVIN	P58J54 bos taurus
7	1411.5	57.4	498	2 Q7ZWH3	Q7zwh3 brachydanio
8	1407	57.3	482	2 Q8AYP6	Q8ayp6 gallus gall
9	1167	47.5	266	2 Q8VDJ4	Q8vdj4 mus musculu
10	1131	46.0	248	2 Q8WZ05	Q8wz05 homo sapien
11	1125.5	45.8	246	2 Q8SP51	Q8sp51 ovib aries
12	948	38.6	507	1 GTR6 HUMAN	Q9ugq3 homo sapien
13	946	38.5	497	2 Q6BTN2	Q6btn2 m mus muscu
14	928	37.8	465	2 Q6GJ34	Q6gj34 xenopus lae
15	894	36.4	445	2 Q8NCC2	Q8nc2 homo sapien
16	810.5	33.0	497	2 Q7Q12	Q7q12 anopheles g
17	804.5	32.7	488	2 Q8MKK4	Q8mkk4 drosophila
18	789.5	32.1	506	2 Q8P59	Q8p59 drosophila
19	788.5	32.1	450	2 Q7PIR5	Q7pir5 anopheles g
20	788.5	32.1	489	2 Q9V609	Q9v609 drosophila
21	788.5	32.1	857	2 Q9V608	Q9v608 drosophila
22	788	32.1	539	2 Q9VU17	Q9vul17 drosophila
23	787	32.0	471	2 Q8RQ46	Q8rq46 drosophila
24	764.5	31.1	433	2 Q9V610	Q9v610 drosophila
25	756	30.8	480	2 Q7PCM5	Q7pcm5 anopheles g
26	756	30.8	483	2 Q7PCH1	Q7pch1 anopheles g
27	736	30.0	487	2 Q93YP9	Q93yp9 arabisopsis
28	734	29.9	488	2 Q9FRL3	Q9frl3 arabisopsis
29	727	29.6	465	2 Q9V848	Q9v848 drosophila
30	724.5	29.5	515	2 Q9LN48	Q9ln48 arabisopsis
31	724	29.5	501	2 Q688U6	Q688u6 oryza sativ

32	723	29.4	501	2	Q9XHW3	Q9xhw3 oryza sativ
33	722	29.4	490	2	Q39416	Q39416 beta vulgar
34	710.5	28.9	496	2	Q8T0T6	Q8t0t6 drosophila
35	702	28.6	501	2	Q688W0	Q688w0 oryza sativ
36	686	27.9	492	2	Q7PR34	Q7pr34 anopheles g
37	684	27.8	482	2	Q8VZ15	Q8vz15 arabisopsis
38	675	27.5	482	2	Q8LBI9	Q8lbi9 arabisopsis
39	670	27.3	463	2	Q8LFR8	Q8lfr8 arabisopsis
40	666.5	27.1	444	2	Q7KQX0	Q7kqx0 drosophila
41	662	26.9	463	2	Q9ZU87	Q9zu87 arabisopsis
42	653.5	26.6	463	2	Q9MAA4	Q9maa4 arabisopsis
43	650.5	26.5	480	2	Q9LEG2	Q9leg2 lycopersico
44	647	26.3	491	2	Q9VNJ2	Q9vnj2 drosophila
45	638	26.0	529	2	Q7QJU9	Q7qju9 anopheles g

ALIGNMENTS

RESULT 1					
Q8WU29	PRELIMINARY;	PRT;	477	AA.	
AC	Q8WU29				
DT	01-MAR-2002 (Tremblrel. 20, Created)				
DT	01-MAR-2002 (Tremblrel. 20, Last sequence update)				
DT	01-MAR-2004 (Tremblrel. 26, Last annotation update)				
DE	Solute carrier family 2, (Facilitated glucose transporter) member				
DE	8.				
GN	Name=SLC2A8;				
OS	Homo sapiens (Human).				
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;				
OC	Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.				
OX	NCBI_TaxID=9606;				
RN	[1]_TaxID=9606;				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Skin;				
RX	MEDLINE=2238257; PubMed=12477932; DOI=10.1073/pnas.242603899;				
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,				
RA	Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,				
RA	Altshuler S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,				
RA	Hopkins R.F., Jordan H., Moore I., Max S.I., Wang J., Hsieh F.,				
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,				
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,				
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,				
RA	Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,				
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,				
RA	Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,				
RA	Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,				
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,				
RA	Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,				
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,				
RA	Krzywinski M.I., Skalska U., Smailus D.E., Scherch A., Schein J.E.,				
RA	Jones S.J., Marra M.A.;				
RT	"Generation and initial analysis of more than 15,000 full-length human				
RT	and mouse cDNA sequences."				
RL	Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).				
RN	[2]				
RP	SEQUENCE FROM N.A.				
RC	TISSUE=Skin;				
RA	Strausberg R.;				
RL	Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.				
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).				
DR	EMBL; BC013043; AAH13043.1; -.				
DR	GO; GO:0016021; C:integral to membrane; IEA.				
DR	GO; GO:0005355; P:glucose transporter activity; IEA.				
DR	GO; GO:0005351; P:sugar porter activity; IEA.				
DR	GO; GO:0005215; P:transporter activity; IEA.				
DR	GO; GO:0008643; P:carbohydrate transport; IEA.				
DR	Pfam; PF00083; Sugar Cr; 1.				
DR	PRINTS; PR00172; GLUCRNSPORT.				
DR	PRINTS; PR00171; SUGTRNSPORT.				
DR	TIGRFAMs; TIGR00879; SP; 1.				

DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 477 AA; 50819 MW; 0B480F94AF063316 CRC64;

Query Match 100.0%; Score 2457; DB 2; Length 477;
Best Local Similarity 100.0%; Pred. NO. 8.8e-164; Indels 0; Gaps 0;
Matches 477; Conservative 0; Mismatches 0;

QY 1 MTPEDDEETQPLLGPPGSGAPRRGVFLAAFAAALGPLSGFALGYSPALPSLQRAAPP 60
DB 1 MTPEDDEETQPLLGPPGSGAPRRGVFLAAFAAALGPLSGFALGYSPALPSLQRAAPP 60

QY 61 APRLDAAASWFGAVVTLGAAAGGVLGWLVDRAKRLSLCSVPFVAGFAVITAAQDV 120
DB 61 APRLDAAASWFGAVVTLGAAAGGVLGWLVDRAKRLSLCSVPFVAGFAVITAAQDV 120

QY 121 WMLGGRLTGLACGVASLAVPVYISETIAYPAVRGLGSCVQLMVVVGILLAYLAGWLE 180
DB 121 WMLGGRLTGLACGVASLAVPVYISETIAYPAVRGLGSCVQLMVVVGILLAYLAGWLE 180

QY 181 WRWLAVLGVCPVPSLMLLMCFMPEPTPRFLITQRRQEAALRFLWGSEQGWEDPPIGAE 240
DB 181 WRWLAVLGVCPVPSLMLLMCFMPEPTPRFLITQRRQEAALRFLWGSEQGWEDPPIGAE 240

QY 241 QSFHLALLRQPGIYKFFIIGVSLMAFQQLSGVNAVFAETIFEEAKFKDSSLASVVG 300
DB 241 QSFHLALLRQPGIYKFFIIGVSLMAFQQLSGVNAVFAETIFEEAKFKDSSLASVVG 300

QY 301 IQVLFTAVALLMDRAGRRLILVLSGVNVFSTSAFGVFKLTGGPGNSSHVATSA 360
DB 301 IQVLFTAVALLMDRAGRRLILVLSGVNVFSTSAFGVFKLTGGPGNSSHVATSA 360

QY 361 AQPVDASVGLAVLAVGSMCLFIAGFAVGMPPIWLMSEIFFLHVKGVATGICVLTNWLM 420
DB 361 AQPVDASVGLAVLAVGSMCLFIAGFAVGMPPIWLMSEIFFLHVKGVATGICVLTNWLM 420

QY 421 AFLVTFEFSLSMELVLPYGAFWLASAFCLPSVLFTLCVPETKGLTLEQITAHFGR 477
DB 421 AFLVTFEFSLSMELVLPYGAFWLASAFCLPSVLFTLCVPETKGLTLEQITAHFGR 477

RESULT 2

GT88_HUMAN STANDARD; PRT; 477 AA.
AC Q9NY64; Q9NSC4;
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 8
DE (Glucose transporter type 8) (Glucose transporter type XI).
GN Name=SLC2A8; Synonyms=GLUT8, GLUTX1;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
OX NCBI_TaxID=9606;
[1]
SEQUENCE FROM N.A.
RP TISSUE=Testis;
RC MEDLINE=20283667; PubMed=10821868; DOI=10.1074/jbc.275.21.16275;
RX Dege H., Schuermann A., Bahrenberg G., Brauers A., Joost H.-G.;
RA "GLUT8, a novel member of the sugar transport facilitator family with
RT glucose transport activity";
RL J. Biol. Chem. 275:16275-16280 (2000).
[2]
SEQUENCE FROM N.A.
RP MEDLINE=10671487; DOI=10.1074/jbc.275.7.4607;
RX Ibberson M.R., Uldry M.A., Thorens B.;
RA GLUTX1, a novel mammalian glucose transporter expressed in the
RT central nervous system and insulin-sensitive tissues";
RL J. Biol. Chem. 275:4607-4612 (2000).
CC -1- FUNCTION: Insulin-regulated facilitative glucose transporter.

Binds cytochalasin B in a glucose-inhibitable manner. Seems to be a dual-specific sugar transporter as it is inhibitable by fructose (By similarity).
-1- SUBCELLULAR LOCATION: Integral membrane protein. Principally intracellular. May move between intracellular vesicles and the plasma membrane. The dileucine internalization motif is critical for intracellular sequestration (By similarity).
-1- TISSUE SPECIFICITY: Highly expressed in testis, but not in testicular carcinoma. Lower amounts present in most other tissues.
-1- INDUCTION: In testis, down-regulated by estrogen.
-1- SIMILARITY: Belongs to the sugar transporter family. Glucose transporter subfamily.

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

EMBL; Y17801; CAB89809.1; --
EMBL; AJ245937; CAB75702.1; --
Genew; HGNC:13812; SLC2A8.
H-InvDB; HIX0008396; --
MIM; 605245; --
GO; GO:0005887; C:integral to plasma membrane; TAS.
GO; GO:0005355; F:glucose transporter activity; TAS.
GO; GO:0005975; P:carbohydrate metabolism; TAS.
GO; GO:0015758; P:glucose transport; TAS.
InterPro; IPR000803; Gluc_transporter.
InterPro; IPR007114; MFS.
InterPro; IPR005828; Sub_transporter.
InterPro; IPR005829; Sugar_transporter.
InterPro; IPR003663; Sugar_transp.
Pfam; PF00083; Sugar_tr; 1.
PRINTS; PR00172; GLUCTRNSPORT.
PRINTS; PR00171; SUGTRNSPORT.
TIGRFAMs; TIGR00879; SP; 1.
PROSITE; PS00850; MFS; 1.
PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Glycoprotein; Multigene family; Sugar transport; Transmembrane; Transport.
FT DOMAIN 1 25 Cytoplasmic (Potential).
FT TRANSMEM 26 46 1 (Potential).
FT DOMAIN 47 70 Extracellular (Potential).
FT TRANSMEM 71 91 2 (Potential).
FT DOMAIN 92 96 Cytoplasmic (Potential).
FT TRANSMEM 97 117 3 (Potential).
FT DOMAIN 118 127 Extracellular (Potential).
FT TRANSMEM 128 148 4 (Potential).
FT DOMAIN 149 156 Cytoplasmic (Potential).
FT TRANSMEM 157 177 5 (Potential).
FT DOMAIN 178 182 Extracellular (Potential).
FT TRANSMEM 183 203 6 (Potential).
FT DOMAIN 204 256 Cytoplasmic (Potential).
FT TRANSMEM 257 277 7 (Potential).
FT DOMAIN 278 292 Extracellular (Potential).
FT TRANSMEM 293 313 8 (Potential).
FT DOMAIN 314 319 Cytoplasmic (Potential).
FT TRANSMEM 320 340 9 (Potential).
FT DOMAIN 341 367 Extracellular (Potential).
FT TRANSMEM 368 388 10 (Potential).
FT DOMAIN 389 404 Cytoplasmic (Potential).
FT TRANSMEM 405 425 11 (Potential).
FT DOMAIN 426 438 Extracellular (Potential).
FT TRANSMEM 439 459 12 (Potential).
FT DOMAIN 460 477 Dileucine internalization motif (By similarity).
FT SITE 12 13 N-linked (GlcNAc...) (By similarity).
FT CARBOHYD 349 349 S -> N (in Ref. 2).
FT CONFLICT 377 377

```
FT CONFLICT 456 457 FS -> LP (in Ref. 2).
SQ CONFLICT 462 462 T -> I (in Ref. 2).
SQ SEQUENCE 477 AA; 50792 MW; 0B480F94B40AEE76 CRC64;

Query Match
Best Local Similarity 99.5%; Score 2445; DB 1; Length 477;
Matches 475; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 MTPEDPEETQPLGGPGGAPRRRVFLAFAAALGFLSGFALGYSPPAIPSLQRAAPP 60
DB 1 MTPEDPEETQPLGGPGGAPRRRVFLAFAAALGFLSGFALGYSPPAIPSLQRAAPP 60

QY 61 APRLDDAAASWFGAVVTLGAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDV 120
DB 61 APRLDDAAASWFGAVVTLGAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDV 120

QY 121 WMLLGGRLITGLACGVASIVAPYIIEIAYPAVRGLLGSVCQLMVVVGILLAYLAGWVLE 180
DB 121 WMLLGGRLITGLACGVASIVAPYIIEIAYPAVRGLLGSVCQLMVVVGILLAYLAGWVLE 180

QY 181 WRWLAVLGCVPVPSMLLLMCFMPTPRFLITQHRROEAMAAALRFLNGSQGWEDPPIGAE 240
DB 181 WRWLAVLGCVPVPSMLLLMCFMPTPRFLITQHRROEAMAAALRFLNGSQGWEDPPIGAE 240

QY 241 QSFHLALLRQPGIYKPFIIIGVSLMAFOQLSGVNAVMPYAEITPEEAKPKDSSLASVVVG 300
DB 241 QSFHLALLRQPGIYKPFIIIGVSLMAFOQLSGVNAVMPYAEITPEEAKPKDSSLASVVVG 300

QY 301 IQVLFTVAALINDRAGRRLLLVLSGVVMVFSFSAFGAYFKLTQGGPGNSSHVAISAPV 360
DB 301 IQVLFTVAALINDRAGRRLLLVLSGVVMVFSFSAFGAYFKLTQGGPGNSSHVAISAPV 360

QY 361 AQVDASVGLAWLAVGSMCLFIAGPAVGWGPPIWLLMSEIFPLHVGKVGATGICVLNNWLM 420
DB 361 AQVDASVGLAWLAVGSMCLFIAGPAVGWGPPIWLLMSEIFPLHVGKVGATGICVLNNWLM 420

QY 421 AFLVTKFESSLMELVRPYGAFWLASAFICFVSLFTFCVPETKGTLEQITAHFEGR 477
DB 421 AFLVTKFESSLMELVRPYGAFWLASAFICFVSLFTFCVPETKGTLEQITAHFEGR 477

RESULT 3
Q6XUI2 PRELIMINARY; PRT; 478 AA.
AC Q6XUI2;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE Glucose transporter 8.
GN Names=GLUT8;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RA Zhao F.-Q., Dong B., Zheng Y.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -1- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AY208940; AAP43920.1; -.
DR GO; GO:0018021; C:integral to membrane; IEA.
DR GO; GO:0005351; P:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transpt.
DR Pfam; PF00083; Sugar_tr; 1.
DR PRINTS; PR00171; SUGRTRANSPORT.
```

```
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
DR SUGAR_TRANSPORT; Transmembrane; Transport.
SQ SEQUENCE 478 AA; 51416 MW; CBE71DAABCAC4CA CRC64;

Query Match
Best Local Similarity 90.8%; Score 2230.5; DB 2; Length 478;
Matches 425; Conservative 26; Mismatches 26; Indels 1; Gaps 1;

QY 1 MTPEDPEETQPLGGPGGAPRRRVFLAFAAALGFLSGFALGYSPPAIPSLQRAAPP 60
DB 1 MTPEDPEETQPLGGPGGAPRRRVFLAFAAALGFLSGFALGYSPPAIPSLQRAAPP 60

QY 61 APRLDDAAASWFGAVVTLGAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDV 120
DB 61 APRLDDAAASWFGAVVTLGAAAGVGLGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDV 120

QY 121 WMLLGGRLITGLACGVASIVAPYIIEIAYPAVRGLLGSVCQLMVVVGILLAYLAGWVLE 180
DB 121 WMLLGGRLITGLACGVASIVAPYIIEIAYPAVRGLLGSVCQLMVVVGILLAYLAGWVLE 180

QY 181 WRWLAVLGCVPVPSMLLLMCFMPTPRFLITQHRROEAMAAALRFLNGSQGWEDPPIGAE 240
DB 181 WRWLAVLGCVPVPSMLLLMCFMPTPRFLITQHRROEAMAAALRFLNGSQGWEDPPIGAE 240

QY 241 -QSFHLALLRQPGIYKPFIIIGVSLMAFOQLSGVNAVMPYAEITPEEAKPKDSSLASVVVG 299
DB 241 HQDFHVAQURRPGVYKPFIIIGVSLMAFOQLSGVNAVMPYAEITPEEAKPKDSSLASVVVG 300

QY 300 VTQVLFTVAALINDRAGRRLLLVLSGVVMVFSFSAFGAYFKLTQGGPGNSSHVAISAPV 359
DB 301 VTQVLFTVAALINDRAGRRLLLVLSGVVMVFSFSAFGAYFKLTQGGPGNSSHVAISAPV 360

QY 360 SAQPDVASVGLAWLAVGSMCLFIAGPAVGWGPPIWLLMSEIFPLHVGKVGATGICVLNNWLM 419
DB 361 SMEADTNGVLAWLAVGSMCLFIAGPAVGWGPPIWLLMSEIFPLHVGKVGATGICVLNNWLM 420

QY 420 AFLVTKFESSLMELVRPYGAFWLASAFICFVSLFTFCVPETKGTLEQITAHFEGR 477
DB 421 MAFLVTKFESSLMELVRPYGAFWLASAFICFVGLFTLACVLPETKGTLEQITAHFEGR 478

RESULT 4
GTR8_RAT STANDARD; PRT; 478 AA.
AC Q9JUZ1; Q9JMA6;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 8
DE (Glucose transporter type 8) (Glucose transporter type X1).
GN Name=Slc2a8; Synonyms=Glut8, Glutx1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A., CHARACTERIZATION, AND MUTAGENESIS OF
RP 12-LEU-LEU-13.
RC TISSUE=Testis;
RX MEDLINE=20138191; PubMed=10671487; DOI=10.1074/jbc.275.7.4607;
RA Ibberson M.R., Udry M.A., Thorens B.;
RT "GLUTX1, a novel mammalian glucose transporter expressed in the
RL central nervous system and insulin-sensitive tissues.";
RN J. Biol. Chem. 275:4607-4612(2000).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RA Ishibashi K.;
RT "Molecular cloning of a new putative glucose transporter.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
```

[3] TISSUE SPECIFICITY.
 RX MEDLINE=20283667; PubMed=10821868; DOI=10.1074/jbc.275.21.16275;
 RA Doege H., Schuermann A., Bahrenberg C., Brauers A., Joost H.-G.;
 RT "GLUT8, a novel member of the sugar transport facilitator family with
 RT glucose transport activity";
 RL J. Biol. Chem. 275:16275-16280(2000).
 CC -!- FUNCTION: Insulin-regulated facilitative glucose transporter.
 CC Binds cytochalasin B in a glucose-inhibitable manner. Seems to be
 CC a dual-specific sugar transporter as it is inhibitable by
 CC fructose.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Principally
 CC intracellular. May move between intracellular vesicles and the
 CC plasma membrane. The dileucine internalization motif is critical
 CC for intracellular sequestration.
 CC -!- TISSUE SPECIFICITY: Highly expressed in adult and pubertal testis,
 CC but not prepubertal testis. Moderate expression in hypothalamus,
 CC cerebellum, brainstem, hippocampus, and adrenal gland. Lower
 CC amounts present in most other tissues.
 CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
 CC transporter subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement. (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AJ245985; CAB75729.1; -;
 CC EMBL; AB033418; BAA94383.1; -;
 CC InterPro; IPR000803; Gluc_transporter.
 CC InterPro; IPR007114; MFS.
 CC InterPro; IPR005828; Sub_transporter.
 CC InterPro; IPR005829; Sugar_transporter.
 CC InterPro; IPR003663; Sugar_transp.
 CC Pfam; PF00083; Sugar_tr_1.
 CC PRINTS; PR00172; GLUCTRNSPRT.
 CC PRINTS; PR00171; SUGTRNSPRT.
 CC TIGRfam; TIGR00879; SP; 1.
 CC PROSITE; PS00850; MFS; 1.
 CC PROSITE; PS00216; SUGAR_TRANSPORT_1; 2.
 CC PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 CC Glycoprotein; Multigene family; Sugar transport; Transmembrane;
 KW Transport.
 FT DOMAIN 1 25 Cytoplasmic (Potential).
 FT TRANSMEM 26 46 1 (Potential).
 FT DOMAIN 47 70 Extracellular (Potential).
 FT TRANSMEM 71 91 2 (Potential).
 FT DOMAIN 92 96 Cytoplasmic (Potential).
 FT TRANSMEM 97 117 3 (Potential).
 FT DOMAIN 118 127 Extracellular (Potential).
 FT TRANSMEM 128 148 4 (Potential).
 FT DOMAIN 149 156 Cytoplasmic (Potential).
 FT TRANSMEM 157 177 5 (Potential).
 FT DOMAIN 178 182 Extracellular (Potential).
 FT TRANSMEM 183 203 6 (Potential).
 FT DOMAIN 204 257 Cytoplasmic (Potential).
 FT TRANSMEM 258 278 7 (Potential).
 FT DOMAIN 279 293 Extracellular (Potential).
 FT TRANSMEM 294 314 8 (Potential).
 FT DOMAIN 315 320 Cytoplasmic (Potential).
 FT TRANSMEM 321 341 9 (Potential).
 FT DOMAIN 342 368 Extracellular (Potential).
 FT TRANSMEM 369 389 10 (Potential).
 FT DOMAIN 390 405 Cytoplasmic (Potential).
 FT TRANSMEM 406 426 11 (Potential).
 FT DOMAIN 427 439 Extracellular (Potential).
 FT TRANSMEM 440 460 12 (Potential).
 FT DOMAIN 461 478 Cytoplasmic (Potential).
 FT SITE 12 13 Dileucine internalization motif.
 FT CARBOHYD 350 N-linked (GlcNAc...) (Probable).

FT MUTAGEN 12 13 LL->AA: Changes subcellular location
 FT mainly to the plasma membrane, thereby
 FT increasing transport activity.
 FT CONFLICT 83 84 GG -> QGA (in Ref. 2).
 SQ SEQUENCE 478 AA; 51458 MW; 95841FC1F18C9E9 CRC64;
 Query Match 88.0%; Score 2162.5; DB 1; Length 478;
 Best Local Similarity 85.4%; Pred. No. 3.5e-143;
 Matches 408; Conservative 33; Mismatches 36; Indels 1; Gaps 1;
 QY 1 MTPEDPTEQLPGPGGAPRGRRVFLAAFAAALGSLSGFALGYSPAIPLQRAAPP 60
 DB 1 MSPEDPTEQLPGPGGAPRGRRVFLAAFAAALGSLSGFALGYSPAIPLQRAAPP 60
 QY 61 APRLDAAAGWFGAVVTGLGAAAGVGLGWLVDRAKRLSLLLCSVPVAVGFAVITAARDV 120
 DB 61 ALRLGDTAAWFGAVVTGLGAAAGVGLGWLVDRAKRLSLLLCSVPVAVGFAVITAARDV 120
 QY 121 WMLGGRLLTGLACGVASLVAPVYISIAYPVAVRGLGSCVQLMVMVVGILLAYVAGWVLE 180
 DB 121 WMLGGRLLTGLACGVASLVAPVYISIAYPVAVRGLGSCVQLMVMVVGILLAYVAGWVLE 180
 QY 181 WRNLAVLCVPPSLMILLMCFMPTPRFLITQHRQEMAAALRELWGSEGWEDPPIGAE 240
 DB 181 WRNLAVLCVPPSLMILLMCFMPTPRFLITQHRQEMAAALRELWGSEGWEDPPIGAE 240
 QY 241 -QSFHLALLRPGIYKPIIGVSLMAFQQLSGVNAVVFYAEIPEEAKFKDSSLASVVVG 299
 DB 241 HQGQALMLRRPGVHKPLIIGICLMVFQQLSGVNAVVFYAEIPEEAKFKDSSLASVVVG 300
 QY 300 VIQVLTAVAALIMDRAGRRLLLVLGVMVVFSAFGAYPKLTQGGFGNSHVAISAPV 359
 DB 301 IIQVLTAVAALIMDRAGRRLLLVLGVMVVFSAFGAYPKLTQGGFGNSHVAISAPV 360
 QY 360 SAQPVASVGLAWLAVGSMCLFIAGFVAGVGPFWLMSFPLHVGATGICVLNTWL 419
 DB 361 SAEPAVHLGLAWLAVGSMCLFIAGFVAGVGPFWLMSFPLHVGATGICVLNTWL 420
 QY 420 MAFLVTKEFSSLMELRYPYGAFWLASAFIPSVLFTLFCVPETKGTLEQITAHFEGR 477
 DB 421 MAFLVTKEFSSLMELRYPYGAFWLASAFIPSVLFTLFCVPETKGTLEQITAHFEGR 478
 RESULT 5
 GTR8_MOUSE STANDARD; PRT; 477 AA.
 AC Q9JIF3; Q9JJP4; Q9JUZ0;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DE 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 8
 DE (Glucose transporter type 8) (Glucose transporter type X1).
 GN Name=SLC2a8; Synonyms=GLUT8, GlutX1;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20138191; PubMed=10671487; DOI=10.1074/jbc.275.7.4607;
 RA Ibberson M.R., Uldry M.A., Thorens B.;
 RT "GLUTX1, a novel mammalian glucose transporter expressed in the
 RT central nervous system and insulin-sensitive tissues";
 RL J. Biol. Chem. 275:4607-4612(2000).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Testis;
 RX MEDLINE=20283667; PubMed=10821868; DOI=10.1074/jbc.275.21.16275;
 RA Doege H., Schuermann A., Bahrenberg C., Brauers A., Joost H.-G.;
 RT "GLUT8, a novel member of the sugar transport facilitator family with
 RT glucose transport activity";
 RL J. Biol. Chem. 275:16275-16280(2000).
 RN [3]

SEQUENCE FROM N.A.
 RC STRAIN=129; TISSUE=Embryonic carcinoma;
 RX MEDLINE=20319023; PubMed=10860996; DOI=10.1073/pnas.97.13.7313;
 RA Carayannopoulos M.O., Chi M.M.-Y., Cui Y., Pingsterhaus J.M.,
 RA McKnight R.A., Mueckler M., Devaskar S.U., Moley K.H.;
 RT "GLUT8 is a glucose transporter responsible for insulin-stimulated
 RT glucose uptake in the blastocyst.";
 RL Proc. Natl. Acad. Sci. U.S.A. 97:7313-7318(2000).
 RN [4]
 RP SEQUENCE FROM N.A.
 RC STRAIN=129/Ola; TISSUE=Spleen;
 RX MEDLINE=21547794; PubMed=11689004; DOI=10.1006/bbrc.2001.5866;
 RA Schepers A., Doege H., Joost H.-G., Schuermann A.;
 RT "Mouse GLUT8: genomic organization and regulation of expression in
 RT 3T3-L1 adipocytes by glucose.";
 RL Biochem. Biophys. Res. Commun. 288:969-974(2001).
 RC -!- FUNCTION: Insulin-regulated facilitative glucose transporter.
 CC Binds cytochalasin B in a glucose-inhibitable manner. Seems to be
 CC a dual-specific sugar transporter as it is inhibitable by
 CC fructose.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Principally
 CC intracellular. May move between intracellular vesicles and the
 CC plasma membrane. The dileucine internalization motif is critical
 CC for intracellular sequestration (By similarity). Insulin induces a
 CC change in the intracellular localization and gives rise to
 CC insertion in the plasma membrane.
 CC -!- TISSUE SPECIFICITY: Highest level of expression in placenta and
 CC testis. Highly expressed in adult and pubertal testis, but not
 CC prepubertal testis. Lower levels of expression in brain, liver,
 CC heart, kidney, fat and skeletal muscle.
 CC -!- DEVELOPMENTAL STAGE: High expression in blastocysts.
 CC -!- INDUCTION: Inhibited under glucose deprivation.
 CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
 CC transporter subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AJ245936; CAB75719.1; -;
 DR EMBL; Y17802; CAB89815.1; -;
 DR EMBL; AF232061; AAF78366.1; -;
 DR EMBL; AJ413951; CAC89690.1; -;
 DR MGD; MGI:1860103; Slc2a8.
 DR GO; GO:0005887; C:integral to plasma membrane; IDA.
 DR GO; GO:0005536; F:glucose binding; IDA.
 DR GO; GO:0005355; F:glucose transporter activity; IDA.
 DR GO; GO:0015758; P:glucose transport; IDA.
 DR GO; GO:0008286; P:insulin receptor signaling pathway; IDA.
 DR GO; GO:0001666; P:response to hypoxia; IDA.
 DR InterPro; IPR000803; Gluc transporter.
 DR InterPro; IPR007114; MFS.
 DR InterPro; IPR005828; Sub transporter.
 DR InterPro; IPR005829; Sug transporter.
 DR InterPro; IPR003663; Sugar_transp.
 DR Pfam; PF00083; Sugar_tr; 1.
 DR PRINTS; PR00172; GLUCTRNSPORT.
 DR PRINTS; PR00171; SUGRTRNSPORT.
 DR TIGRFAMs; TIGR00879; SP; 1.
 DR PROSITE; PS00850; MFS; 1.
 DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
 DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
 KW Glycoprotein; Multigene family; Sugar transport; Transmembrane;
 KW Transport.
 FT DOMAIN 1 25 Cytoplasmic (Potential).
 FT DOMAIN 26 46 1 (Potential).
 FT DOMAIN 47 70 Extracellular (Potential).
 FT DOMAIN 71 91 2 (Potential).
 FT DOMAIN 92 96 Cytoplasmic (Potential).

FT TRANSMEM 97 117 3 (Potential).
 FT DOMAIN 118 127 Extracellular (Potential).
 FT TRANSMEM 128 148 4 (Potential).
 FT DOMAIN 149 156 Cytoplasmic (Potential).
 FT TRANSMEM 157 177 5 (Potential).
 FT DOMAIN 178 182 Extracellular (Potential).
 FT TRANSMEM 183 203 6 (Potential).
 FT DOMAIN 204 257 Cytoplasmic (Potential).
 FT TRANSMEM 258 278 7 (Potential).
 FT DOMAIN 279 293 Extracellular (Potential).
 FT TRANSMEM 294 314 8 (Potential).
 FT DOMAIN 315 320 Cytoplasmic (Potential).
 FT TRANSMEM 321 341 9 (Potential).
 FT DOMAIN 342 367 Extracellular (Potential).
 FT TRANSMEM 368 388 10 (Potential).
 FT DOMAIN 389 404 Cytoplasmic (Potential).
 FT TRANSMEM 405 425 11 (Potential).
 FT DOMAIN 426 438 Extracellular (Potential).
 FT TRANSMEM 439 459 12 (Potential).
 FT DOMAIN 460 477 Cytoplasmic (Potential).
 FT SITE 12 13 Dileucine internalization motif (By similarity).
 FT CARBOHYD 350 350 N-linked (GlcNAc...) (By similarity).
 FT CONFLICT 39 39 S -> N (in Ref. 1).
 FT CONFLICT 94 94 S -> A (in Ref. 2 and 4).
 FT CONFLICT 429 429 S -> N (in Ref. 1).
 SQ SEQUENCE 477 AA; 51523 MW; A3753PB34B452F9A CRC64;
 Query Match 87.7%; Score 2156; DB 1; Length 477;
 Best Local Similarity 86.0%; Pred. No. 1e-142;
 Matches 411; Conservative 31; Mismatches 34; Indels 2; Gaps 2;
 QY 1 MTPEDPEETQPLLLGPPGGSAPGRVFLAAFAAALGPLSGFALGYSSPAISLQRAAPP 60
 DB 1 MSPEPDQETQPLRLRPPERTPRGRVFLASFAAALGPLSGFALGYSSPAISLRTAPP 60
 QY 61 APRLDDAASWFGAVVTGAAAGVGLGWLVDRAGRKLSLLCSVPFVFAFVITAQDV 120
 DB 61 ALRLGDNAASWFGAVVTGAAAGVGLGWLDRSGRKLSSLLCTVPFVTFGAVITAARDV 120
 QY 121 WMLLGGRLITGLACGVASIVAPVYISETAYPAVRGLGSCVQLMVVVGILLAYLAGWVLE 180
 DB 121 WMLLGGRLITGLACGVASIVAPVYISETAYPAVRGLGSCVQLMVVVGILLAYLAGWVLE 180
 QY 181 WRWLAVLGCVPPLSMLLMCMFETPRFLTLTQRRQEAAMALRFLWGSQGWEDPPIGAE 240
 DB 181 WRWLAVLGCVPPLTMLLMCMYMPETPRFLTLTQRRQEAAMALRFLWGSQGWEDPPIGAE 240
 QY 241 -QSFLHALLRQPGIYKPIIIGVSLMAFQOLSGVNAMFYAETIFFEAKFKDSSLASVVVG 299
 DB 241 HQGFQALLRRPGIYKPLIIGISLMVFQOLSGVNAMFYANIFFEAKFKDSSLASVVVG 300
 QY 300 VIQVLTAVALLIMDRAGRRLLLVSGVWVVESTAFGNAYFKLTQGGPGNSSHVAISAPV 359
 DB 301 IIQVLTAVALLIMDRAGRRLLLVSGVWVVESTAFGNAYFKLTQGGPGNSSHVGL-VPI 359
 QY 360 SAQPYDASVGLAWLAVGSMCLFIAGVWGPIPLMLLSEIFPLHVKGVATGCVLTNNL 419
 DB 360 AAEFVDVQVGLAWLAVGSMCLFIAGVWGPIPLMLLSEIFPLHVKGVATGCVLTNNF 419
 QY 420 MAFLVTKFSSLMVELRYPGAFWLASAFICFISVLFTLCVPTKTKGLTQIHTAHFGR 477
 DB 420 MAFLVTKFSSWMELRYPGAFWLTAFAFCALSVLFTLTVPVTKTKGLTQIHTAHFGR 477
 RESULT 6
 ID GTR8 BOVIN STANDARD; PRT; 334 AA.
 AC P58354;
 DT 28-FEB-2003 (Rel. 41, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Solute carrier family 2, facilitated glucose transporter, member 8

DE (Glucose transporter type 8) (Glucose transporter type X1) (Fragment).
GN Name=SLC2A8; Synonym=GLUT8, GLUTX1;
OS Bos taurus (Bovine).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Bovinae; Bos.
OX NCBI_TaxID=9913;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21482567; PubMed=11599048; DOI=10.1002/mrd.1099;
RA Augustin R., Pocar P., Navarrete-Santos A., Wrenzycki C., Gandolfi F.,
RA Niemann H., Fischer B.;
RT "Glucose transporter expression is developmentally regulated in in
RT vitro derived bovine preimplantation embryos.";
RL Mol. Reprod. Dev. 60:370-376(2001).
CC -!- FUNCTION: Insulin-regulated facilitative glucose transporter.
CC Binds cytochalasin B in a glucose-inhibitable manner. Seems to be
CC a dual-specific sugar transporter as it is inhibitable by fructose
CC (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Principally
CC intracellular. May move between intracellular vesicles and the
CC plasma membrane. The di-leucine internalization motif is critical
CC for intracellular sequestration (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
CC transporter subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See [http://www.isb-sib.ch/](http://www.isb-sib.ch/announce/)
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AF321324; AAK69606.1; -;
DR InterPro; IPR000803; Gluc transporter.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00083; Sugar cr; 1.
DR PRINTS; P00172; GLUCTRNSPORT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00217; SUGAR TRANSPORT 1; 1.
DR PROSITE; PS00217; SUGAR TRANSPORT 2; PARTIAL.
KW Multigene family; Sugar transport; Transmembrane; Transport.
FT NON_TER 1
FT DOMAIN 1 14 Cytoplasmic (Potential).
FT TRANSMEM 15 35 5 (Potential).
FT DOMAIN 36 38 Extracellular (Potential).
FT TRANSMEM 39 59 6 (Potential).
FT DOMAIN 60 113 Cytoplasmic (Potential).
FT TRANSMEM 114 134 7 (Potential).
FT DOMAIN 135 149 Extracellular (Potential).
FT TRANSMEM 150 170 8 (Potential).
FT DOMAIN 171 176 Cytoplasmic (Potential).
FT TRANSMEM 177 197 9 (Potential).
FT DOMAIN 198 224 Extracellular (Potential).
FT TRANSMEM 225 245 10 (Potential).
FT DOMAIN 246 261 Cytoplasmic (Potential).
FT TRANSMEM 262 282 11 (Potential).
FT DOMAIN 283 295 Extracellular (Potential).
FT TRANSMEM 296 316 12 (Potential).
FT DOMAIN 317 334 Cytoplasmic (Potential).
FT SEQUENCE 334 AA; 36699 MW; 0EE9B670ADAB71DD CRC64;
Query Match 62.6%; Score 1537.5; DB 1; Length 334;
Best Local Similarity 87.7%; Pred. No. 1.3e-99;
Matches 293; Conservative 17; Mismatches 23; Indels 1; Gaps 1;
QY 145 ISFIAPVAVGLGSCVQLMVMVVGILLAYLAGVLEWRMLAVLGCVPSPSLMLLLMCFMPE 204
DB 1 ISFIAPVAVGLGSCVQLMVMVVGILLAYLAGVLEWRMLAVLGCVPSPSLMLLLMCFMPE 60

QY 205 TPPELLTQHRQRQAMALRELWGSEQWEDPPICAE-QSFHLALLRQPIKPIIGVSL 263
DB 61 TPPELLSQHKHQEAMAAQFLMGVAQGWEEPLGAQHQDFHVAQLRPGVYKPIIGISL 120
QY 264 MAFQOLSGVNAVMPYATITFEFAKFKDSSLASVVGVIQVLFTAVAILINDRAGRRLILV 323
DB 121 MAFQOLSGVNAVMPYATITFEFAKFKDSSLASVVGVIQVLFTATAALINDRAGRRLILT 180
QY 324 LSGVVMVFSTSAFGAYFKLTQGGPGNSSHVAISAPVSAQPDASVGLAWLAVGSMCLFIA 383
DB 181 LSGVVMVFSTSAFGTYFKLTQGGPGNSSHVDLPALVNEADTNVGLAWLAVGNNCLFIA 240
QY 384 GFVAVGWPPIWLLMSEIFPLHVKGVAATGICVLTNWMAFLVTKFSSLMELVLPYGAFWL 443
DB 241 GFVAVGWPPIWLLMSEIFPLHVKGVAATGVCVLTNWMAFLVTKFSSLMELVLPYGAFWL 300
QY 444 ASAFICFISVLFTLPCVPETKGTLEQITAHFEGR 477
DB 301 ASAFICFISVLFTLPCVPETKGTLEQITAHFEGR 334
RESULT 7
QY2WH3
ID Q7ZWH3 PRELIMINARY; PRT; 498 AA.
AC Q7ZWH3;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Solute carrier family 2, (Facilitated glucose transporter), member
DE 8.
GN ORFNames=zgc:56364;
OS Brachydanio rerio (Zebrafish) (Danio rerio).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
OC Cyprinidae; Danio.
OC NCBI_TaxID=7955;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole body;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.D., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Haieh F.,
RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalls D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences".
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Whole body;
RA Strausberg R.;
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; BC049409; AHA49409.1; -;
DR ZFIN; ZDB-GENE-030829-25; zgc:56364.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005335; F:glucose transporter activity; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.

Db 240 SVLFTLTVPVETKGRLEQVTAHFEGR 266

RESULT 10
Q8WZ05 PRELIMINARY; PRT; 248 AA.

ID Q8WZ05;
AC AC
DT 01-WAR-2002 (TrEMBLrel. 20, Created)
DT 01-WAR-2002 (TrEMBLrel. 20, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Zhang P.P., Zhou X.M., Jiang H.Q., Huang Y., Qin W.X., Zhao X.T.,
RA Wan D.F., Gu J.R.;
RC Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RL -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR EMBL; AP289587; AAL5571.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0003555; F:glucose transporter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR Pfam; PF00083; Sugar tr; 1.
DR PRINTS; PR00172; GLUCTRNSPORT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Hypothetical protein; Transmembrane.
SQ SEQUENCE 248 AA; 27017 MW; AA929A749BBEE58 CRC64;

Query Match 46.0%; Score 1131; DB 2; Length 248;
Best Local Similarity 93.0%; Pred.No. 2.6e-71;
Matches 227; Conservative 5; Mismatches 6; Indels 6; Gaps 2;

Qy 164 MVVVGILLAYLAGWVLWRWLAVLGCVPSLMILLMCFMPETPRFLLTQHRROEAMALR 223
Db 1 MVVVGILLAYLAGWVLWRWLAVLGCVPSLMILLMCFMPETPRFLLTQHRROEAMALR 60
Qy 224 FLWGSEQGWEDPPIGAEQSFHLALLQPGLYKPFIIIGVSLMAFQQLSGVNAVVFYAETIF 283
Db 61 FLWGSEQGWEDPPIGAEQSFHLALLQPGLYKPFIIIGVSLMAFQQLSGVNAVVFYAETIF 120
Qy 284 EEAKPKDSSLASVVVGVIQVLFATAALMDRAGRRLILLVSGVMVFTSAFGAYFKLT 343
Db 121 EEAKPKDSSLASVVVGVIQVLFATAALMDRAGRRLILLVSGVMVFTSAFGAYFKLT 180
Qy 344 OQGPGNSSHVAIAPVSQAOPVDASVGLAWLAVGSMCLFIAGFAVGWGP-IPWLLMSIIF 402
Db 181 OQGPGNSSHVAIAPVSQAOPVDASVGLAWLAVGSMCLFIAG-----GFQALWSLLACLRF 235
Qy 403 LHKV 406
Db 236 LHK 239

RESULT 11
Q8SPS1 PRELIMINARY; PRT; 246 AA.

ID Q8SPS1;
AC AC
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Glucose transporter 8 (Fragment).
OS Eukarya; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
OC Caprinae; Ovis.
OX NCBI_TaxID=9940;
RN [1]
RP SEQUENCE FROM N.A.

RC TISSUE=Placenta;
RX PubMed=15013641; DOI=10.1016/j.placenta.2003.08.012;
RA Limesand S.W., Regnault T.R.H., Hay W.W. Jr.;
RT "Characterization of glucose transporter 8 (GLUT8) in the ovine
RT placenta of normal and growth restricted fetuses.";
RL Placenta 25:70-77(2004).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AF495799; AAM18513.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005351; F:sugar porter activity; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0008443; P:carbohydrate transport; IEA.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR003663; Sugar transport.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00083; Sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
KW Sugar transport; Transmembrane; Transport.
FT NON TER 246 246
SQ SEQUENCE 246 AA; 26537 MW; E8F3871946CD0723 CRC64;
Query Match 45.8%; Score 1125.5; DB 2; Length 246;
Best Local Similarity 88.2%; Pred. No. 6.4e-71;
Matches 217; Conservative 11; Mismatches 17; Indels 1; Gaps 1;
QY 219 MAALRLNSEQWEDPPPTGAE-QSHALLRPPGYKPIIIGVSLMAFQQLSGVNAVMP 277
DB 1 MAAMQFLWSAGPWEPPPTGAHQGFHVAQLRCPGYKPIIIGISLMAFQQLSGVNAVMP 60
QY 278 YAEITFEAKFKDSSLASVVGVIOVLTAVAILMDRAGRLLLVLSGVVWVFTSARG 337
DB 61 YAEITFEAKFKDSSLASVVGVIOVLTAVAILMDRAGRLLLVLSGVVWVFTSARG 120
QY 338 AYFKLFGGGPNSHVAISAPVASQPDASVGLAVLSMCLFIAGFAVGWGPPIPWLLM 397
DB 121 AYFKLFGGGPNSHVDLPAPVSMEPANTNVGLAVLSMCLFIAGFAVGWGPPIPWLLM 180
QY 398 SEIFPLHVGKATGICVLTNLWMAFLVTFKFSLSMEVLPYGFNWLASAFICFSVLTFLF 457
DB 181 SEIFPLHVGKATGVCVLTNWFNMAFLVTFKFSLSMEVLPYGFNWLASAFICFVLTFLA 240
QY 458 CVPETK 463
DB 241 CVPETK 246
RESULT 12
GTR6_HUMAN
ID GTR6_HUMAN STANDARD; PRT; 507 AA.
AC Q9UGG3;
DT 28-FEB-2003 (Rel. 41, Created)
DT 28-FEB-2003 (Rel. 41, Last sequence update)
DE 05-JUL-2004 (Rel. 44, Last annotation update)
DE Solute carrier family 2, facilitated glucose transporter, member 6
DE (Glucose transporter type 6) (Glucose transporter type 9).
GN Names=SLC2A6; Synonyms=GLUT9;
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Leukocyte;
RX MEDLINE=20427701; PubMed=10970791; DOI=10.1042/0264-6021:3500771;
RA Doeghe H., Bocianski A., Joost H.-G., Schuermann A.;
RT "Activity and genomic organization of human glucose transporter 9
RT (GLUT9), a novel member of the family of sugar-transport facilitators
RT predominantly expressed in brain and leucocytes.";
RL Biochem. J. 350:771-776(2000).

RNA SEQUENCE FROM N.A.
RP TISSUE=Lymphoid;
RA Young J.M., Woodward K.J., Aziz S., Burley M., Kwiatkowski D.J.,
RA Povey S.;
RT "Cloning of a sugar transporter gene, a G-beta subunit like gene and
RT three novel genes in human chromosome 9q34.";
RL Submitted (SEP-1998) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RP TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udwin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Scherch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Facilitative glucose transporter; binds cytochalasin B
CC with low affinity.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. The dileucine
CC internalization motif is critical for intracellular sequestration
CC (By similarity).
CC -!- TISSUE SPECIFICITY: Highly expressed in brain, spleen and
CC peripheral blood leukocytes.
CC -!- SIMILARITY: Belongs to the sugar transporter family. Glucose
CC transporter subfamily.
CC -!- CAUTION: Has been described as GLUT9 in literature, but this gene
CC name has already been used for SLC2A9.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).

CC EMBL; Y17803; CAB96996.1; -.
DR EMBL; AJ011372; CAB66155.1; -.
DR EMBL; BC013740; AAH13740.1; -.
DR Genew; HGNC:11011; SLC2A6.
DR H-InvDB; HIX0008509; -.
DR MIM; 606813; -.
DR InterPro; IPR000803; Gluc_transporter.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub_transporter.
DR InterPro; IPR005829; Sug transporter.
DR InterPro; IPR003663; Sugar transport.
DR Pfam; PF00083; Sugar tr; 1.
DR PRINTS; PR00172; GLUCTRNSPORT.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Glycoprotein; Multigene family; Sugar transport; Transmembrane;
KW Transport.
FT DOMAIN 1 37 Cytoplasmic (Potential).
FT

```
FT TRANSMEM 38 58 1 (Potential).
FT DOMAIN 59 81 Extracellular (Potential).
FT TRANSMEM 82 102 2 (Potential).
FT DOMAIN 103 111 Cytoplasmic (Potential).
FT TRANSMEM 112 132 3 (Potential).
FT DOMAIN 133 140 Extracellular (Potential).
FT TRANSMEM 141 161 4 (Potential).
FT DOMAIN 162 168 Cytoplasmic (Potential).
FT TRANSMEM 169 189 5 (Potential).
FT DOMAIN 190 194 Extracellular (Potential).
FT TRANSMEM 195 215 6 (Potential).
FT DOMAIN 216 289 Cytoplasmic (Potential).
FT TRANSMEM 290 310 7 (Potential).
FT DOMAIN 311 314 Extracellular (Potential).
FT TRANSMEM 315 335 8 (Potential).
FT DOMAIN 336 339 Cytoplasmic (Potential).
FT TRANSMEM 340 360 9 (Potential).
FT DOMAIN 361 395 Extracellular (Potential).
FT TRANSMEM 396 416 10 (Potential).
FT DOMAIN 417 435 Cytoplasmic (Potential).
FT TRANSMEM 436 456 11 (Potential).
FT DOMAIN 457 482 Extracellular (Potential).
FT TRANSMEM 483 483 Cytoplasmic (Potential).
FT DOMAIN 484 507 Dileucine internalization motif (By similarity).
FT SITE 5 6 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 370 370 M -> T (in Ref. 2).
FT CONFLICT 500 500
FT SITE 507 AA; 54569 MW; 1FC6E2EB11588460 CRC64;

Query Match
Best Local Similarity 38.6%; Score 948; DB 1; Length 507;
Matches 216; Conservative 80; Mismatches 166; Indels 44; Gaps 11;

QY 10 QPLG-----PPGSAPGR-----RVFLAAFAAALGPLSPFGALGYSPGA 50
DB 3 EFLLAGGPDYDFTPEKPPSPGDRARVGTQKRVFLATFAVLGNFSFGVALVTSVP 62
QY 51 IFSLOAAAPAPRLDAAASWFGAVVTLLGAAGGVLGGWLVDRAGKLSLLCSVPFVAG 110
DB 63 IFALERSLDPDHLTKSQASWFGSVFTLLGAAGGUSAMLNDLLGRKUSIMFSAVPSAAG 122
QY 111 FAVITAAQDVMMLGRLTLGLACGVASIVAPVYISEIAYPAVRGLGSCVQLVMVVGIL 170
DB 123 YALMAGAHGLMMLLGRUTLGTGAGGLTRACIPVYVSEIAPPVGRVGLGATPQLMAVFGSL 182
QY 171 LAYLAGVLEWLVAVLGCVPSPSMLLMCFMPETPRFLTOHRRQEAALRFLWGSEQ 230
DB 183 SLVALGLLPMRWLAVAGEAPVLMILLSPFNPSRFLSRGRDEEALRALWLRGTDV 242
QY 231 G--WEDPPIG---AEQSFLH--ALLRQGVYKPFITIGVSLMAFOOLSGVNAVWFVAETIF 283
DB 243 DVHWEFEQIDNRRQSSVSWAERAPVHCPIITVALLMRLLQQTGTLPTLVYLQSFIF 302
QY 284 EBAKF-----KDSLASVVGVIQVLTAVAAALIMDRGRRLLLVLSGVVMVFSTSAFGAY 339
DB 303 DSTAVLLPKDD--AAIVGAVRLLSVLIAALTMDLGRKVLFFVSAIMFAANLTGLY 359
QY 340 FKLTGGGPNSHVAISAPVS-----AQPVDASVG-LAWIAGVGNCLFIAGFVAGWGP 392
DB 360 IHF---GPRPLSPNLTAGLESWGDLAQPLAAPAGYLTLPVLLATMLFIMGYAVGWGPI 416
QY 393 PMLLMSEIFPLHVKGVATGICVLTWLMALVTKFSSIMELVRLPYGAFWLASAFICFSV 452
DB 417 TWLLMSEVLPARGVAGSLCVLASWLTAFVLTKSFLPVSTFGVLQVFPFFFAALCLVSL 476
QY 453 LPTLCVETPKGKTLEQITAFHE-GR 477
DB 477 VFTGCCVETPKGRSLEQIESFRMGR 502
```

RESULT 13
Q8BTN2

```
ID AC Q8BTN2 PRELIMINARY; PRT; 497 AA.
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-MAR-2004 (TRENBLrel. 26, Last annotation update)
DE Mus musculus NOD-derived CD11c +ve dendritic cells cDNA, RIKEN full-
DE length enriched library, clone:F630103L12 product:similar to SOLUTE
DE CARRIER FAMILY 2, FACILITATED GLUCOSE TRANSPORTER, MEMBER 6 (GLUCOSE
DE TRANSPORTER TYPE 6) (GLUCOSE TRANSPORTER TYPE 9).
GN Name=Sic2a6;
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=99279253; PubMed=10349636; DOI=10.1016/S0076-6879(99)03004-9;
RA Carninci P., Hayashizaki Y.;
RT "High-efficiency full-length cDNA cloning.";
RL Meth. Enzymol. 303:19-44(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=21085660; PubMed=11217851; DOI=10.1038/35055500;
RA RIKEN FANTOM Consortium;
RT "Functional annotation of a full-length mouse cDNA collection.";
RL Nature 409:685-690(2001).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=20499374; PubMed=11042159; DOI=10.1101/gr.145100;
RA Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT "Normalization and subtraction of cap-trapper-selected cDNAs to
RT prepare full-length cDNA libraries for rapid discovery of new genes.";
RL Genome Res. 10:1617-1630(2000).
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RX MEDLINE=20530913; PubMed=11076861; DOI=10.1101/gr.152600;
RA Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA Konno H., Akiyama J., Nishi K., Kitsuai T., Tashiro H., Itoh M.,
RA Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA Fujiwaki S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsura S., Kawai J.,
RA Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT "RIKEN integrated sequence analysis (RISA) system-384-format
RT sequencing pipeline with 384 multicapillary sequencer.";
RL Genome Res. 10:1757-1771(2000).
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=NOD;
RA Adachi J., Aizawa K., Akimura T., Arakawa T., Bono H., Carninci P.,
RA Fukuda S., Furuno M., Hanagaki T., Hara A., Hashizume W.,
RA Hayashida K., Hayatsu N., Hiramoto K., Hirakawa T., Hirozane T.,
RA Hori F., Imotani K., Ishii Y., Itoh M., Kagawa I., Kasukawa T.,
RA Kato H., Kawai J., Kojima Y., Kondo S., Konno H., Kouda M., Koya S.,
RA Kurihara C., Matsuyama T., Miyazaki A., Murata M., Nakamura M.,
RA Nishi K., Nomura K., Numazaki R., Ohno M., Ohsato N., Okazaki Y.,
RA Saito R., Saitoh H., Sakai C., Sakai K., Sakazume N., Sano H.,
RA Sasaki D., Shibata K., Shinagawa A., Shiraki T., Sogabe Y., Tagami M.,
RA Tagawa A., Takahashi F., Takaku-Akaira S., Takeda Y., Tanaka T.,
RA Tomaru A., Toya T., Yaeunishi A., Muramatsu M., Hayashizaki Y.;
```


Submitted (Apr-2002) to the EMBL/GenBank/DBJ databases.
-!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
-!- SIMILARITY: Belongs to the sugar transporter family.
DR EMBL; AK089246; BAC40811.1; -.
DR MGI; MGI:2443286; Slc2a6.
DR CO; GO:0016020; C-membrane; ISS.
DR GO; GO:0005355; F:glucose transporter activity; ISS.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005828; Sub transporter.
DR InterPro; IPR003663; Sugar transp.
DR InterPro; IPR005829; Sug transporter.
DR Pfam; PF00883; Sugar tr; 1.
DR PRINTS; PR00171; SUGTRNSPORT.
DR TIGRFAMs; TIGR00879; SP; 1.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transport; Transmembrane; Transport.
SQ SEQUENCE 497 AA; 54445 MW; 8E14D1DCD4B244483 CRC64;

Query Match 38.5%; Score 946; DB 2; Length 497;
Best Local Similarity 43.6%; Pred. No. 4.6e-58;
Matches 214; Conservative 77; Mismatches 168; Indels 32; Gaps 9;

QY 3 PEDPETQLPGPP-----GSAPGRVRVFLAFAAALGFLSFGFALGYSSPAIPSLQRAA 58
Db 17 PEVP-----ATPGERERAGALKNRRVFLATFAAVLGNFSGYALVYTSPIPELKLSS 69
QY 59 PAPRLDDAAASWFGAVVTLGAAGGVGLWLVDRAGKLSLLLCSPVPFAGFAVITAQA 118
Db 70 DPALHLDKTOASWFGSVFTLGAAGLSAMLLNDLLGRKLSIMFSAVPSAIGYAINAGAR 129
QY 119 DVWMLLGGRLTLGLACCVASLVAPVYISIAVPAVRGLGSCVQLMNVVGVILLAYLAGWV 178
Db 130 GLWMLLGLRMLTGFGAGLTAACTPVYVSEIAPDVKGALGATPQLMAVFGSLVALGLL 189
QY 179 LEWRNLAVLCVPPSLMLLLCMFPTPRFLTLQHRROEAMALRFLWG-SQGWEDPPI 237
Db 190 LPWRNLAVAGEPVLITMILLSPNSPRFLSKSRDEALQALTLWRADSEVHWEFEQI 249
QY 238 G---AEQSFLH--ALLRQGIYKPFILIGVSLMAFQOLSGNVMFVAETIPEAK-FKDS 291
Db 250 QDNVRQSSGRVSWAEAREPRVYRPLVIAVLMRFLOQLTGITPILVYLQTFIDNTSVLPS 309
QY 292 SLASVVVGVQLVFTAAALIMDRAGRLLLVLSGVVVFSTSAFGAYFK-----LTQGG 346
Db 310 QQDAALVGVRLLSVLIAAVTMDLAGRKVLLVYSVMFAANLTGLYVQFVPRPLT--- 366
QY 347 PGNSSHVAISAPVSAQPVDSVGLAWLAVGSMCLFIAGFVAGVGPPIPLMLSEIPLHYK 406
Db 367 PNSTVEIVLGD-----AFNYLTLPILLATMLFMGVAMGWPITLLMSEVLPLRAR 420
QY 407 GVATGICVLTNMLAFVLTKEPSSLMNEVLRPGAFWLASAFCLFSVLFTLCVPETKGT 466
Db 421 GVASGLCVLVSMLTAFVLTNLYFLAVNAPGLQVPPFFPSAICLLSLLTGCCVPETGRS 480
QY 467 LEQITAHFEGR 477
Db 481 LEQIEAFFTR 491

RESULT 14

Q66J34 PRELIMINARY; PRT; 465 AA.
AC Q66J34; PRELIMINARY;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE MGC82056 protein.
GN Name=MGC82056;
OS Xenopus laevis (African clawed frog).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;

Query Match 37.8%; Score 928; DB 2; Length 465;
Best Local Similarity 41.9%; Pred. No. 7.8e-57;
Matches 198; Conservative 90; Mismatches 152; Indels 32; Gaps 6;

QY 26 VFLEAFAAALGFLSFGFALGYSSPAIPSLQRAAPPAPRLDDAA-----ASWFGAVVTL 78
Db 1 MFLEAFAVAVLGNFTFGYALVYTSPIPALEK-----DDQLHINAEISWFGAVFAL 52
QY 79 GAAAGVVGWLVDRAGKLSLLLCSPVPFAGFAVITAQDVWMLLGGRLTLGLACGVAS 138
Db 53 GACAGGIGSMFLNDRLGRKLSIMFSAVPSLSGLFLMGSQAHSMLLGLRILTFGAGWMTS 112
QY 139 LVAPYVISEIAYPAVRGLLGSVCVQLMNVVGVILLAYLAGWLVSRWLAVLVCVPPSLMLLL 198
Db 113 SSIPVYIIEISHGVGGGLGACPCQIMAVCGSLVILVALGULLPWRMLAAGEVTVTMLLL 172
QY 199 MCFMPTPRFLTLQHRROEAMALRFLMGSEQWEDPPIGABQSFLALLR----- 249

```

Db 173 LCFMPSDFRLIAKGDKALALWLGANTDYQ-----GEVERIKSNILKSSSTLSWTE 228
QY 250 --QPGYKPIICVSLMAFQQLSGVNAVVFYAEIPEEAK-KFQSSLASVVVGVQVLF 306
Db 229 LQOPYYYKILLIAVFRFLQQLSGVSPILYLETIFNRTKVLIRGGYDAALVGVRLLSV 288
QY 307 AVAALIMDRAGRLILVLGVVVFSTSAFGAYFKLTQGGPGNSSHVAISAPVSAQVDA 366
Db 289 IISASVMDKAGRKILYTSSTLMFVSSLSMGLYVHTVDINHNSNTRTMSISSAEPSEP 348
QY 367 SVGLAWLVAGSCLEFIAGFAVGKPIPLMLMSEIPLPHVKGVATGICVLTNMLMAFLVTK 426
Db 349 VNYIQLILICIMLYIYAFGKGPITWLLMSEILPKSRGVASGLCVVSVHIAFILTE 408
QY 427 EFSSLMELVRPFGAFWLASAFICFVSLFTFCVPETKGTLEQITAHFE-GR 477
Db 409 AFIPVVNTLSLOTPEFFFTAVCAASIMFTYFFVFPETKGTLEQIESYFRTGR 460

```

RESULT 15

```

Q8NCC2
ID Q8NCC2 PRELIMINARY; PRT; 445 AA.
AC Q8NCC2;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein FLJ90355.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
NCBI_TaxID=9606;
RN [1]_
SEQUENCE FROM N.A.
RA Isozaki T., Ota T., Nishikawa T., Hayaashi K., Otsuki T., Sugiyama T.,
RA Suzuki Y., Nagai K., Sugano S., Ishii S., Kawai-Hio Y., Saito K.,
RA Yamamoto J., Wakamatsu A., Nakamura Y., Kojima S., Nagahari K.,
RA Masuho Y., Ono T., Okano K., Yoshikawa Y., Aotsuka S., Sasaki N.,
RA Hattori A., Okumura K., Iwayanagi T., Ninomiya K.;
RL Submitted (MAR-2002) to the EMBL/GenBank/DBSJ databases.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005355; F:glucose transporter activity; IEA.
DR GO; GO:0005351; P:sugar porter activity; IEA.
DR GO; GO:0005215; P:transporter activity; IEA.
DR GO; GO:0008643; P:carbohydrate transport; IEA.
DR InterPro; IPR000803; Gluc_transporter.
DR InterPro; IPR007114; MFS.
DR InterPro; IPR005028; Sub_transporter.
DR InterPro; IPR003663; Sugar_transpt.
DR InterPro; IPR005829; Sug_transporter.
DR Pfam; PF00083; Sugar_cr; 1.
DR PRINTS; PR00172; GLUCTRNSPORT.
DR PRINTS; PR00171; SUGTRNSPORT.
DR PROSITE; PS00850; MFS; 1.
DR PROSITE; PS00216; SUGAR_TRANSPORT_1; 1.
DR PROSITE; PS00217; SUGAR_TRANSPORT_2; 1.
KW Sugar transporter; Transmembrane; Transporter.
SQ SEQUENCE 445 AA; 48040 MW; BF37DC0C313A32CE CRC64;

```

```

Query Match 36.4%; Score 894; DB 2; Length 445;
Best Local Similarity 40.1%; Pred. No. 1.8e-54;
Matches 200; Conservative 73; Mismatches 134; Indels 92; Gaps 9;
QY 10 QPLIG-----PPGSGAPGR-----RVFLAAFAALGPLSGFALGYSSPA 50
Db 3 EPLLGAEGFDYDTFFPEKPPSPGDRARVGTQLQNKRVFLATFAAVLGNFSGVALVYTSVP 62
QY 51 IPSLQRAAPPAPRLDDAAAFGAVVTIGAAAGVVGWLVDRAGRKLSLLLCSPVPFVAG 110
Db 63 IPALERSLDPLHLTKSQASWFSVFTLGAAGAGGLSAMLNDLLGRKLSIMPSAVPSAAG 122

```

```

QY 111 FAVITAAQDVMMLLGRRLLTGLACGVASIVAPVYISETAYPAVRGLGSCVQLMVVVGIL 170
Db 123 YALMAGAHGLWMLLGRLLTGTGAGGLTAACIPVYVYSEIAPPVGRGALGATPQLMAVFGSL 182
QY 171 LAYLAGWLEWRWLAIVLGCVPSPSLMLLMCFMPEPFRLLTOHRRQEAALRFLWGSQ 230
Db 183 SLIYALGLLLPWKWLAVAGEAPVLINILLSPFNSPRLLSRGRDEEALRALAWLRGTDV 242
QY 231 G--WEDPPIG---ABQSPHL--ALLRQPIGYKPIIIGVSLMAFQQLSGVNAVVFYAEI 283
Db 243 DVHWEFEQIQDNRVRSQSSRVSWAEARPHVCKRPITVALLMRLQLQTGITPLVYLQSF 302
QY 284 BEAKF----KQSSLASVVVGVIOVLTFAVALINDRAGRRLLLVLSGVVMPSTSAFGAY 339
Db 303 DSTAVLLPPKOD---AAIVGAVRLLSVLIAALTMDLAGRKVL----- 341
QY 340 FKLTQGGPGNSSHVAISAPVSAQPVDAVGLAWLAVGSMCLFIAGFAYGVWGPIPWLLMSE 399
Db 342 -----LFVSGYAVGVWGPIPWLLMSE 361
QY 400 IFPLHVKVATGICVLTNMLMAFLVTKFSSLMELVRPFGAFWLASAFICFVSLFTFCV 459
Db 362 VLPLRARGVASGLCVLASWLTAFVLTKSFLPVVSTFGVLQVDFPFFFAALCLVSLVFTGCCV 421
QY 460 PETKGTLEQITAHFE-GR 477
Db 422 PETKGRSLEQIESPFRTRGR 440

```

Search completed: February 24, 2005, 05:52:57
Job time : 183 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 25, 2005, 16:32:02 ; Search time 5101 Seconds
(without alignments)
4531.099 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457
Sequence: 1 MTPEDPFTQPLLPGGSA.....CVPETKGTLEQITAHFEGR 477

Scoring table: BLOSUM62 Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09886954/runat_23022005_154242_13208/app_query.fasta_1.647
-DB=GenEmbl -QFMT=fastap -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdd -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09886954 @CGN 1.1 5600 @runat_23022005_154242_13208 -NCFU=6 -ICPU=3
-NO MWAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : GenEmbl.*

1: gb.ba.*
2: gb.htg.*
3: gb.in.*
4: gb.om.*
5: gb.ov.*
6: gb.pat.*
7: gb.ph.*
8: gb.pl.*
9: gb.pr.*
10: gb.ro.*
11: gb.sts.*
12: gb.sy.*
13: gb.un.*
14: gb.vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2457	100.0	2080	6 AX179740	AX179740 Sequence
2	2457	100.0	2145	9 BC019043	BC019043 Homo sapi
3	2448	99.6	1873	9 HSA245937	AJ245937 Homo sapi
4	2448	99.6	2217	6 AX076667	AX076667 Sequence

5	2445	99.5	1508	9 HSA17801	Y17801 Homo sapien
6	2230.5	90.8	2073	4 AY208940	AY208940 Bos tauru
7	2162.5	88.0	2087	6 AX076669	AX076669 Sequence
8	2162.5	88.0	2087	10 RNO245935	AJ245935 Rattus no
9	2159	87.9	1490	10 MMU17802	Y17802 Mus musculu
10	2156	87.7	1795	10 AF232061	AF232061 Mus muscu
11	2150	87.5	2072	6 AX076671	AX076671 Sequence
12	2150	87.5	2072	10 MMU245936	AJ245936 Mus muscu
13	2146	87.3	2189	10 AB033418	AB033418 Rattus no
14	1554.5	63.3	1461	6 AX191507	AX191507 Sequence
15	1337.5	62.6	1012	4 AF321324	AF321324 Bos tauru
16	1411.5	57.4	1929	5 BC049409	BC049409 Danio rer
17	1407	57.3	1449	5 AB083371	AB083371 Gallus ga
18	1268	51.6	789	6 AX191497	AX191497 Sequence
19	1167	47.5	1493	10 BC021758	BC021758 Mus muscu
20	1125.5	45.8	738	4 AF495799	AF495799 Ovis arie
21	980.5	39.9	668	6 CQ723656	CQ723656 Sequence
22	972.5	39.6	2011	6 AX076679	AX076679 Sequence
23	958.5	39.0	1580	9 HSA17803	Y17803 Homo sapien
24	958.5	39.0	2356	6 AX677165	AX677165 Sequence
25	958.5	39.0	2356	9 BC013740	BC013740 Homo sapi
26	958.5	39.0	2390	6 CQ782893	CQ782893 Sequence
27	958.5	39.0	2390	6 BD127367	BD127367 Primer fo
28	958.5	39.0	2390	9 AK074927	AK074927 Homo sapi
29	958.5	39.0	2487	9 HSA011372	AJ011372 Homo sapi
30	958.5	39.0	2696	6 CQ412537	CQ412537 Sequence
31	949	38.6	1541	6 AX076677	AX076677 Sequence
32	945.5	38.5	1661	6 CQ723535	CQ723535 Sequence
33	934	38.0	3694	5 BC081076	BC081076 Xenopus l
34	910.5	37.1	2364	10 BC058210	BC058210 Mus muscu
35	904.5	36.8	2171	6 CQ782786	CQ782786 Sequence
36	904.5	36.8	2171	6 BD127313	BD127313 Primer fo
37	904.5	36.8	2171	9 AK074836	AK074836 Homo sapi
38	891.5	36.3	9582	10 MMU413951	AJ413951 Mus muscu
39	891.5	36.3	198006	2 AL731852	AL731852 Mus muscu
40	870.5	35.4	225370	9 AL445222	AL445222 Human DNA
41	867.5	35.3	3114	9 AK096630	AK096630 Homo sapi
C 42	824.5	33.6	250940	2 AC096489	AC096489 Rattus no
C 43	807	32.8	241623	2 AC103008	AC103008 Rattus no
44	804.5	32.7	1412	6 CQ590307	CQ590307 Sequence
45	804.5	32.7	1673	3 AY122079	AY122079 Drosophila

ALIGNMENTS

RESULT 1	AX179740	Sequence 28 from Patent WO0146258.	2080 bp	DNA	linear	PAT 06-AUG-2001
LOCUS	AX179740	Sequence 28 from Patent WO0146258.				
DEFINITION	AX179740					
ACCESSION	AX179740					
VERSION	AX179740.1	GI:15132104				
KEYWORDS						
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE						
AUTHORS	Baughn, M.R., Burford, N., Au-Young, J., Lu, D.A., Yang, J., Reddy, R., Lal, P., Hillman, J.L., Azimzai, Y., Yue, H., Nguyen, D.B., Yao, M.G., Gandhi, A.R., Tang, Y.T. and Khan, F.A.					
TITLE	Transporters and ion channels					
JOURNAL	Patent: WO 0146258-A 28 28-JUN-2001;					
FEATURES	Incyte Genomics, Inc. (US)					
source	Location/Qualifiers					
	1..2080					
	/organism="Homo sapiens"					
	/mol_type="unassigned DNA"					
	/db_xref="taxon:9606"					
	/note="Incyte ID No: 1416107CB1"					

ORIGIN

Alignment Scores:

Pred. No.: 7.58e-167 Length: 2080

[illegible]

REMARK
COMMENT

NIH-MGC Project URL: <http://mgc.nci.nih.gov>
 On Dec 19, 2003 this sequence version replaced gi:17512129.
 Contact: MGC help desk
 Email: gcgaps@mail.nih.gov
 Tissue Procurement: ATCC/DCTD/DRP
 cDNA Library Preparation: Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: National Institutes of Health Intramural
 Sequencing Center (NISC),
 Gaithersburg, Maryland;
 Web site: <http://www.nisc.nih.gov/>
 Contact: nisc_mgc@nhi.nih.gov
 Akhtar, N., Ayelle, K., Beckstrom-Sternberg, S.M., Benjamin, B.,
 Blakesley, R.W., Bouffard, G.G., Green, K., Brinkley, C., Brooks, S.,
 Dietrich, N.L., Granite, S., Guan, X., Gupta, J., Haghighi, P.,
 Hansen, N., Ho, S.-L., Karlins, E., Kwong, P., Laric, P., Legaspi, R.,
 Maduro, Q.L., Masiello, C., Maskeri, B., Mastrian, S.D., McCloskey, J.C.,
 McDowell, J., Pearson, R., Stantripop, S., Thomas, P.J., Touchman, J.W.,
 Tsutgeon, C., Vogt, J.L., Walker, M.A., Wetherby, K.D., Wiggins, L.,
 Young, A., Zhang, L.-H. and Green, E.D.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAL.Plate: 30 Row: n Column: 3
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 21361448.

FEATURES
source

Location/Qualifiers
 1. .2145
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="MGC:20634 IMAGE:4641145"
 /issue_type="Skin, melanotic melanoma."
 /clone_lib="NIH_MGC_20"
 /lab_host="DH10B-R"
 /notes="Vector: pOTB7"
 1. .2145
 /gene="SLC2A8"
 /note="synonyms: GLUTX1, GLUT8"
 /db_xref="LocusID:29988"
 /db_xref="MIM:605245"
 46. .1479
 /gene="SLC2A8"
 /codon_start=1
 /product="solute carrier family 2, (facilitated glucose
 transporter) member 8"
 /protein_id="AAH19043.1"
 /db_xref="GI:17512130"
 /db_xref="LocusID:29988"
 /db_xref="MIM:605245"
 /translation="MTPEDEPTQPLLPGGSPARRRVFLAFAALGPILSFGFAL
 GYSPAPISLQRAAPAPRLDDAASWFGAVVTGLGAAGGVGLVGLVDVDRGRKLSLL
 CSVPFVFAVITAQDVWMLLGRLLTGLACGVASLAPVYISIAYPVAVRGLIGSC
 VOLMVVGLLLAYLAGVLEWRWLAVLVCVPSLLMLLMCFMPTPRFLTLORRQEA
 MAALRLWSEQWEDPPPIQASQFHLALRQPIIGVSLMAFOQLSGVNAV
 FYAETIFEEKFKDLSASVVGVIQVLTAAALMDRAGRLLLVLSGVYVWFST
 ARGAFKLTQGGPGNSHVAISAPVSAQPDASVGLWLAAGSMCLFIAGFVAGWGP
 PWLMSEIFPLHVKGATGICVLTNMLMAFLVTKFSSLMELVRPYGAPFWLASAFICF
 SVLFTLCVPETKGTLEQITAFEGR"

gene

1. .2145
 /gene="SLC2A8"
 /note="synonyms: GLUTX1, GLUT8"
 /db_xref="LocusID:29988"
 /db_xref="MIM:605245"
 46. .1479
 /gene="SLC2A8"
 /codon_start=1
 /product="solute carrier family 2, (facilitated glucose
 transporter) member 8"
 /protein_id="AAH19043.1"
 /db_xref="GI:17512130"
 /db_xref="LocusID:29988"
 /db_xref="MIM:605245"

CDS

1. .2145
 /gene="SLC2A8"
 /note="synonyms: GLUTX1, GLUT8"
 /db_xref="LocusID:29988"
 /db_xref="MIM:605245"
 46. .1479
 /gene="SLC2A8"
 /codon_start=1
 /product="solute carrier family 2, (facilitated glucose
 transporter) member 8"
 /protein_id="AAH19043.1"
 /db_xref="GI:17512130"
 /db_xref="LocusID:29988"
 /db_xref="MIM:605245"

ORIGIN

Alignment Scores:
 Pred. No.: 7,866-167 Length: 2145
 Score: 2457.00 Matches: 477
 Percent Similarity: 100.00% Conservative: 0
 Best Local Similarity: 100.00% Mismatches: 0
 Query Match: 100.00% Indels: 0
 DB: 9 Gaps: 0
 US-09-886-954A-1 (1-477) x BC019043 (1-2145)
 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20

46	ATGACGCCGAGGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGACGCGG	105
21	ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe	40
106	CCCCCGCGCGCGCGCTCTCTCGCGCGCTTTCGCGCGCTGCGCGCGCGCTCAGCTTC	165
41	GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro	60
166	GGCTTTCGCGCTGGCTACAGCTCCCGCGCATCCCTAGCTGCAGCGCGCGCGCGCG	225
61	AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla	80
226	GCCCGCGCTGGAGAGCGCGCTCTCTGGTTCGGGGCTGTCTGAGCCCTGGGTGCC	285
81	AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu	100
286	GGCGCGCGCGGAGTGTCTGGCGCGCTGGCTGGTGGACCGCGCGCGCGCAAGCTGAG	345
101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal	120
346	TTGCTGTGCTCCGTGCGCTTCTGGCGCGCTTTCGCTCATCACCGCGCGCGCGAGCG	405
121	TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal	140
406	TGGATGCTCTGGCGCGCGCTCTCTCACCGCGCTGGCTGGTGTTCCTCCCTAGTG	465
141	AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuGlySerCys	160
466	GCCCGCGCTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGTGTCTCGGCTCTGT	525
161	ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu	180
526	GTCGAGCTAATGTCGTCTGGCATCTCTCTGGCTGCTGGTGTCTGGTGTGGGAG	585
181	TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys	200
586	TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	645
201	PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAla	220
646	TTTCATGCCCGAGACCCCGCTTCTCTGTCTGCTGCTGCTGCTGCTGCTGCTGCTG	705
221	AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProLeuGlyAlaGlu	240
706	GCCCTGGCTTCTGTGGGCTCCGAGCAGGCTGGAGAGAGAGAGAGAGAGAGAGAGAG	765
241	GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleLeuGly	260
766	CAGAGCTTTCACCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG	825
261	ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAlaAlaValMetPheTyrAlaGlu	280
826	GTCTCCCTGATGGCTTTCAGCAGCTGTCGGGGGTCAACCGCGCTCATGTCTTATGAG	885
281	ThrIlePheGluGluAlaLysPheLysSerSerLeuAlaSerValValGlyVal	300
886	ACCATCTTTGAAGAGCGCAAGTTCAAGGACAGCAGCTGGCTGGCTGGCTGGCTGGCT	945
301	IleGlnValLeuPheThrAlaValAlaAlaLeuLeuMetAspArgAlaGlyArgArgLeu	320
946	ATCCAGGTGCTGTTTCAGAGCTGTGGGGCTCTCATCATGGACAGAGCGGCGGAGCTG	1005
321	LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe	340
1006	CTCTCTGGTCTTGTGAGGT	1065
341	LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer	360
1066	AAGCTGAGCCAGGCTGGCGCTGCACTCTCGACGCTGGCGCTCTCGGCTCTCTCTCT	1125
361	AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu	380

Db 1126 GCACAGCCTGTTGATGCCAGCGTGGGGCTGGCTGGCTGGCGTGGCGAGCATGTGSCCTC 1185
Qy 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProLeuLeuMetSerGluile 400
Db 1186 TTATCATGCCGGCTTTGCGGTGGGCTGGGGCCCATCCCTGGCTCCTCATGTGAGAGATC 1245
Qy 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuLeuThrAsnTrpLeuMet 420
Db 1246 TTCCCTCTGCATGTCACGGCGTGGGACAGGCATCTGGCTCCTCACCACTGGCTCATG 1305
Qy 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTrpGlyAla 440
Db 1306 GCCTTTCTCGTACCAAGAGATTACAGAGCTTCATGAGGCTCCTCAGGCCCTATGAGCC 1365
Qy 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1366 TTCTGGCTTGGCTCCGCTTTTCGATCTTCAGTGTCTTTTCACATTTGTCTGTGTCCCT 1425
Qy 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1426 GAACCTNAAGGAAGACTCTGGAACAAATCACAGCCCATTTGAGGGCGGA 1476

RESULT 3
HSA245937
LOCUS HSA245937 1873 bp mRNA linear PRI 18-FEB-2000
DEFINITION Homo sapiens mRNA for glucose transporter (GLUTX1 gene).

ACCESSION AJ245937
VERSION AJ245937.1 GI:7018305
KEYWORDS glucose transporter; GLUTX1 gene.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1
AUTHORS Ibberson, M., Uldry, M. and Thorens, B.
TITLE GLUTX1, a novel mammalian glucose transporter expressed in the
central nervous system and insulin-sensitive tissues
J. Biol. Chem. 275 (7), 4607-4612 (2000)

JOURNAL 20138191
MEDLINE 10671487
PUBMED 2 (bases 1 to 1873)
REFERENCE Ibberson, M.R.
AUTHORS Ibberson, M.R.
TITLE Direct Submission
JOURNAL Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, SWITZERLAND

FEATURES
Location/Qualifiers

1..1873
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"

gene

CDS

1..1873
/gene="GLUTX1"
4..1437
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CAB75702.1"
/db_xref="GI:7018306"
/db_xref="GOA:Q9NY64"
/db_xref="UniProt/Swiss-Prot:Q9NY64"
/translation="WTPDEPTQLPGPGGAPRRRVFLAAFAALGPLSGFAL
GVSSPAISLQRAAPPALPDAAASWFGAVVTLGAAGGVLGWLVDRAGKLSLL
CSVPFAGAVITAQDVWMLLGGRLLTGLACGVASIVAPVYISIAIPAVRGLLGSC
VOLMVVGLLAWGLVLEWRLAVLGCVPSMLLMCFMPETPRLLTQHRKEA
MAALRFLMSEGSDSLASVVGVIQVLFVAALIMDRAGRLLLSGVVMVFNAM
FYAETIFEEKFGKSSVAVIAPVSAQPDASVGLAWGNMCLFVAGVWGPI
PWLMSIPLPLVKGVATGICVLTNWLMAFLVTKFSSLMELVRLPYGAFWLASAFICF
SVLFTLFCVPEIKGKLEQITAHFGR"

ORIGIN

Alignment Scores:

Pred. No.: 2,96e-166 Length: 1873
Score: 2448.00 Matches: 475
Percent Similarity: 99.79% Conservative: 1
Best Local Similarity: 99.58% Mismatches: 1
Query Match: 99.63% Indels: 0
DB: 9 Gaps: 0
US-09-886-954A-1 (1-477) x HSA245937 (1-1873)
Qy 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 4 ATGACGCCCGAGSACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGAGCGCG 63
Qy 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuLeuGlyProLeuSerPhe 40
Db 64 CCCCAGGCGCCCGGGCTTCTCGCGCTTCGCCCTTCGCCCTGGCGGCCCACTCAGCTTC 123
Qy 41 GlyPheAlaLeuGlyTyrosSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db 124 GGCTTGGCGCTCGGCTACAGCTCCCGGCCATCCCTAGCCTGCAGCGCGCGCGCCCG 183
Qy 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db 184 GCCCGCGCGCTGGACGACGCGCGCTCTCTGGTTCGGGGCTGTCTGACCTGGGTGCC 243
Qy 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db 244 CGCGCGGGGGAGTGTGGGGCGGTGGCTGTGTGACCGCGCGCGGCGGCAAGCTGAGCCTC 303
Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 304 TTGTGTGCTCCGTGGCTTCTGGCGCGCTTTCGGCTCATCAGCGCGCGCGCAGACGTG 363
Qy 121 TrpMetLeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 364 TGGAGTGTGTGGGGGGCGGCTCTCTACCGCGCTTGGCTGGGTGTTCCTCCCTAGTG 423
Qy 141 AlaProValTyrlleSerGluIleAlaTyProAlaValArgGlyLeuLeuGlySerCys 160
Db 424 GCCCGGTCTACATCTCCGNAATCGCTACCCAGCAGTCCGGGGGTTCCTCGGCTCCTGT 483
Qy 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyLeuAlaGlyTrpValLeuGlu 180
Db 484 GTGCAGCTAAATGTGTGTGTGGCATCTCTCCGTGCTACCTGGCAGGTGGGTGTGTGAG 543
Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 544 TGGCGCTGGCTGGCTGTGGGTGGCTGGCGTGGCGCTCCCTCATGCTCTCTCATGTGC 603
Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAla 220
Db 604 TTATGCCCGAGACCCCGCGCTTCTGTGCTGACTCAGCACAGCGCGCCAGAGGCCATGGCC 663
Qy 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
Db 664 GCCCTGGGTTCTTGTGGGGTCCGAGCAGGGCTGGGAAGACCCCGCCATCGGGGCTGAG 723
Qy 241 GlnSerPheHisLeuAlaLeuArgGlnProGlyIleTyLeuProPheIleIleGly 260
Db 724 CAGAGCTTTCACCTGGGCGCTGTCTGGCGAGCGCGCGCATCTACAAGCCCTTCATCATCGGC 783
Qy 261 ValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaValMetPheTyAlaGlu 280
Db 784 GTCTCCCTGATGGCTTCCAGCAGCTGTGGGGGTCAACGCCGCTCATGTCTTATGCAGAG 843
Qy 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
Db 844 ACCATCTTTGAAGAGGCCCAAGTTCAAGGACAGCAGCCTTCGGCTCGTGTGGGTGTC 903
Qy 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
Db 904 ATCCAGGTGCTGTTCACAGCTGTGGCGCTCTCATCATGGAACAGAGCGGCGGAGGTG 963

QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 964 CTCCTGCTTGTGTCAGGTGGTTCATGCTTTCAGCAGCAGTGGCTTCGGCGCTACTTC 1023

QY 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
Db 1024 AAGCTAGCCAGGGTGGCCCTGGCAACTCTCGCAGGTGGCCATCTCGGGCTGTCTCT 1083

QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTyrPheLeuAlaValGlySerMetCysLeu 380
Db 1084 GCACAGCCTGTTGATGTCAGCGTGGGCTGGCTGGCTGGCTGGCGCAACATGTGCTC 1143

QY 381 PheIleAlaGlyPheAlaValGlyTyrGlyProIleProTyrPheLeuMetSerGluLe 400
Db 1144 TTCATGCCGCTTTCGGTGGGCTGGGCGCCATCCCTGGCTTCCTCATGTTCAGAGATC 1203

QY 401 PheProLeuHisValGlyValAlaThrGlyLeuCysValLeuThrLeuMet 420
Db 1204 TTCCTCTGATGTCAGGGCGTGGCGACAGGATCTGCTCTCACCACCTGGCTCATG 1263

QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
Db 1264 GCCTTCTGTCAGCAGGATTCAGCAGGCTCATGGAGTCTCATGGCCCTATGAGCC 1323

QY 441 PheTyrLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1324 TTCCTGCTTCCCTGCTTCTGATCTTCAGTGTCTTTCATCTTGTCTGTCCT 1383

QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1384 GAAATTAAGGAAGAACTCTGGAACAAATCACAGCCCATTTTTCAGGGCGGA 1434

RESULT 4
AX076667
LOCUS AX076667 2217 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 1 from Patent WO0104145.
ACCESSION AX076667
VERSION AX076667.1 GI:12711198
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1
Thorens, B., Ibberson, M. and Uldry, M.
TITLE Glutx polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 1 18-JAN-2001;
University of Lausanne (CH)
FEATURES
Location/Qualifiers
1..2217
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
348..1781
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28495.1"
/db_xref="GI:12711199"
/translation="MTPEDPEETQLLPPGGSAPGRBRVFLAFAALGPLSPGFAL
GVSPALPSLQRAAPAPRDLDDNAASFVAVTLGAAGVGLGWLVDRAKRLSILL
CSVPFVAGFAVITAQDQWMLLGGRLTLGLACGASLVAPVYSEIAYPAVRLGLSC
VOLMVVGLLALYAGWLEWRLVAVLGVPPSLMLLMCFMPEPTFRFLTLQRRQEA
MAALRLWSEQGWEDPPICAEOSPHALLRPGIYKPFIIIGVSLMAFQOLSGVNAV
FYAETIFERAKPKDSLVVGVIOVLETAVALIMDRAGRLILVLISGVNVFST
AFGAYFLTGGGPGNSHWALISAPVSAQVFDASVGLAWLAVGNWCLFIAGPAVAGPI
PMLMSEIFPLHWGVATGICVLTNWLMAPLVTKBFSLSLMEVLIRPYGAFWLASAPCFI
SVLFTLFCVPEIKGTLEQITAFHEGR"

ORIGIN
Alignment Scores:
Pred. No.: 3,61e-166 Length: 2217
Score: 2448.00 Matches: 475
Percent Similarity: 99.79% Conservative: 1

Best Local Similarity: 99.58% Mismatches: 1
Query Match: 99.63% Indels: 0
DB: 6 Gaps: 0

US-09-886-954A-1 (1-477) x AX076667 (1-2217)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 348 ATGACGCCCGAGGACCCAGAGGAAACCCAGCGCTTCTGGGGCTCTCTGGCGCAGCGC 407

QY 21 ProArgGlyArgArgValPheLeuAlaIlePheAlaAlaLeuGlyProLeuSerPhe 40
Db 408 CCCCAGCGCGCGCGCTTCTTCCTCGCGCTTTCGGCGCTGCTGGCGCCACCTCAGCTTC 467

QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db 468 GGCCTTGGCGCTGGCTAGCTTCCCGCGCATCTAGCTTGACGGCGCGCGCGCCCG 527

QY 61 AlaProArgLeuAspAspAlaAlaIleSerTyrPheGlyAlaValThrLeuGlyAla 80
Db 528 GCCCGCGCTGGACGACGCGCGCTCTCTGGTTCGGGGCTGCTCGTACCTTGGGTGCC 587

QY 81 AlaAlaGlyGlyValLeuGlyGlyTyrLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db 588 GCGCGCGGGGAGTGTCTGGCGCGCTGGCTGGTGGACCGCGCGCGCGCAAGCTGAGCCTC 647

QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 648 TTGCTGTGCTCGTGGCGCTTCTGGCGCGCTTTCGGCTCATCACCGCGCGCGCGAGAGCTG 707

QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 708 TGGATGCTCTGGGGCGCGCTCTCTCACCGCTGGCTGGGTGTGCTCCCTAGTG 767

QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db 768 GCCCGCGCTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGTGTCTCGGCTCTCTGT 827

QY 161 ValGlnLeuMetValValValGlyLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu 180
Db 828 GTGACGTAAATGGTGTCTGGCATCTCTGGCGCTACCTGGCAGGCTGGGTGTCTGGAG 887

QY 181 TrpArgTyrLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 888 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 947

QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db 948 TTTATGCCCGAGACCCCGCGCTTCTGTGCTGACTCAGCAGCGCGCGCGCGCATGGCC 1007

QY 221 AlaLeuArgPheLeuTyrGlySerGluGlnGlyTyrGluAspProIleGlyAlaGlu 240
Db 1008 GGCCTGGGTTCCTGTGGGGCTTCCGAGCAGGGCTGGGAAGACCCCGCATTCGGGGCTGAG 1067

QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleGly 260
Db 1068 CAGAGCTTTCACCTGGCGCTGTGGCGAGCGCGCATCTACAAGCGCTTTCATCATCGGC 1127

QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
Db 1128 GTCTCCCTGATGCCCTTCAGCAGCTGTCTGGGGTTCAGCGCGCTCATGTTCTATGAGAG 1187

QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
Db 1188 ACATCTTTCAAGAGCCCAAGTTCAAGGACAGCGCTGGCTGGCTGGCTGGCTGGCTGG 1247

QY 301 IleGlnValLeuPheThrAlaValAlaIleMetAspArgAlaGlyArgArgLeu 320
Db 1248 ATCCAGGTGCTGTTCACAGCTGTGGCGCTCTCATCATGACAGCAGCGCGCGAGGCTG 1307

QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 1308 CTCCTGCTTCTGTCAGGTGTGCTCATGTTGTTCAGCAGCAGTGGCTTTCGGCGCTACTTC 1367

QY	341	LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer	360
Db	1368	AAAGTGACCCAGGGTGGCCCTGGCAACTCTCGCACGTGGCCATCTCGCGCGCTGCTCTCT	1427
QY	361	AlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCysLeu	380
Db	1428	GCACAGCCTGTGATGCCAGCGTGGCGCTGGCTGGCTGGCGGCAACATGGCCTC	1487
QY	381	PheIleAlaGlyPheAlaValGlyTyrGlyProIleProTyrLeuLeuMetSerGluile	400
Db	1488	TTTATCCCGCGCTTTCGGGTGGCTGGCGCCATCCCTGCTCTCATGTTCAGAGATC	1547
QY	401	PheProLeuHisValIysGlyValAlaThrGlyIleCysValLeuThrAsnTyrLeuMet	420
Db	1548	TTTCCTCTGCATGTCAAGGCGTGGCGACAGGCATCTGGCTCTCCACCAACTGGCTCATG	1607
QY	421	AlaPheLeuValThrIysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla	440
Db	1608	GCCTTTCTGTGACCAAGAGTTTCAGACGCTCATGGAGTCTCAGGCCCTATGGAGCC	1667
QY	441	PheTyrLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro	460
Db	1668	TTTGTGCTTGCCTCCGCTTCTGCACTTTCAGTGCTCTTTTTCATTTTGTGTGTCTCT	1727
QY	461	GluThrIysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg	477
Db	1728	GAATTAAGGAAGAACTCTGGAACAAATCACAGCCCATTTTGGAGGGCGA	1778
RESULT 5			
HSAL17801			
LOCUS	HSAL17801	1508 bp	mRNA linear PRI 13-MAY-2000
DEFINITION	Homo sapiens mRNA for glucose transporter 8 (GLUT8 gene).		
ACCESSION	Y17801		
VERSION	Y17801.1	GI:7688145	
KEYWORDS	glucose transporter 8; GLUT8 gene.		
SOURCE	Homo sapiens (human)		
ORGANISM	Homo sapiens		
REFERENCE	1 Doerge, H., Schurmann, A., Bahrenberg, G., Brauers, A. and Joost, H.G. GLUT8, a novel member of the sugar transport facilitator family with glucose transport activity J. Biol. Chem. 275 (21), 16275-16280 (2000)		
AUTHORS	J. Biol. Chem. 275 (21), 16275-16280 (2000)		
JOURNAL	20283667		
MEDLINE	10821868		
PUBMED	Joost, H.G.		
REFERENCE	Direct Submission		
AUTHORS	Submitted (16-JUL-1998) H. Joost, Institute of Pharmacology and Toxicology, Technical University Aachen, Wendlingweg 2, D-52057 Aachen, FRG		
JOURNAL	Location/Qualifiers		
FEATURES	1..1508		
source	/organism="Homo sapiens"		
	/mol_type="mRNA"		
	/db_xref="taxon:9606"		
	/tissue_type="testis"		
gene	1..1508		
	/gene="GLUT8"		
CDS	27..1460		
	/gene="GLUT8"		
	/codon_start=1		
	/product="glucose transporter 8"		
	/protein_id="CAB89809.1"		
	/db_xref="GI:7688146"		
	/db_xref="GOA:Q9NY64"		
	/db_xref="UniProt/Swiss-Prot:Q9NY64"		
	/translation="MTPEDEETQPLILGPGSAPRGRRVFLFAAFALGPLGFGFALGYSPALPSLQRAAPPAPRLDAAASFAGAVTLGRAAGVGVLGNLVDRAGRKLSLLCSVPFVAGFAVITTAQDVMLLGRLTLGLACGVASIVAVPISEIAYPAVRGLLSCVLMVVVGIILAYLAGVLEWRWLAIVLCVFPFSLMLLLMLMCFMPEFTRFLDTQHRQEA		

ORIGIN			
Alignment Scores:	3.75e-166	Length: 1508	
Pred. No.:	2445.00	Matches: 475	
Score:	99.58%	Conservative: 0	
Best Local Similarity:	99.58%	Mismatches: 2	
Query Match:	99.51%	Indels: 0	
DB:	9	Gaps: 0	
US-09-886-954A-1 (1-477) x HSAL17801 (1-1508)			
QY	1	MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla	20
Db	27	ATGACGCCCGAGGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGAGCGCG	86
QY	21	ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe	40
Db	87	CCCCGCGCGCGCGCTTCTCTCGCGCTTCTCGCGCTTCTCGCGCTTCTCGCGCTTCT	146
QY	41	GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro	60
Db	147	GGCTTCGGCTTCGGCTACAGCTCCCGGCCATCTTAGCCTGCAGCGCGCGCGCGCG	206
QY	61	AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla	80
Db	207	GCCCGCGCTTGGACGACGCGCGCTTCTCGGTTCGGGGCTGTCTGACCTGGGTGCC	266
QY	81	AlaAlaGlyGlyValLeuGlyGlyTyrLeuValAspArgAlaGlyArgGlyLeuSerLeu	100
Db	267	CGCGCGCGCGGAGTGTCTGGCGGCTTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGT	326
QY	101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal	120
Db	327	TTGTCTGTCTCGTGGCTTCTGTGGCGCTTTCGGCTCATCACCGCGCGCGCGACGTG	386
QY	121	TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal	140
Db	387	TGGATGCTGTCTGGGGCGCGCTCTCTACCGCGCTTGGCTTGGGTGTGTCTCTCTCT	446
QY	141	AlaProValTyrIleSerGluIleAlaValProAlaValArgGlyLeuLeuGlySerCys	160
Db	447	GCCCCGGTCTACATCTCCGAATTCGCTTACCCAGCAGTCCGGGGGTCTCTCGGTCTCT	506
QY	161	ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu	180
Db	507	GTGCAGCTAATGCTGTCTGGCATCTCTCTGGCTTACCTGGCAGGCTGGGTGTGTGAG	566
QY	181	TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys	200
Db	567	TGGCGCTGT	626
QY	201	PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAla	220
Db	627	TTTCATGCCCGAGACCCCGCGCTTCTCTGTGACTCAGACAGCGCGCAGGAGGCTTGC	686
QY	221	AlaLeuArgPheLeuTyrGlySerGluGlnGlyTyrGluAspProIleGlyAlaGlu	240
Db	687	GCCCTGTGGTCTCTGTGGGCTTCCGAGCAGGCGCTGGGAGAGACCCCGCTTCTCAT	746
QY	241	GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGly	260
Db	747	CAGAGCTTTTACCTGCGCTTGTGTGGCGAGCCCGCGCATCTACAGCCCTTCTCAT	806
QY	261	ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu	280
Db	807	GTCTCTCTCATGGCTTTCAGCAGCTGTCTGGGGGTCAACGCGCTCATGTCTTCTAT	866


```
QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyVal 300
Db 867 ACCATCTTTTCAAGAGGCCAAGTTCAAGGACAGCAGCCTGGCTCGTCTCGTGGGTGC 926
QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
Db 927 ATCCAGGTGCTGTTACAGCTGTGGCGCTCTCATCATGACACAGACGAGCGGAGGCTG 986
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 987 CTCCTGGTCTGTGAGTGTGGTCAATGTTTCAGCACAGTGGCTTCGGCGCTACTTC 1046
QY 341 LysLeuThrGlnGlyGlyProGlyAenSerSerHisValAlaLysSerAlaProValSer 360
Db 1047 AAGCTGACCCAGGCTGGCCCTCTGCACTCTGCACTGGCCCTCTGGCGCTGTCTCT 1106
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
Db 1107 GCAACAGCTGTGATGCCAGCGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1166
QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProIleProIleProIleProIle 400
Db 1167 TTTCTGCGCGCTTTCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1226
QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAenTrpLeuMet 420
Db 1227 TTTCCCTCTGATGTCAGGCGCTGGCGACAGGCTCTGGCTCTGCTCTGCTCTGCT 1286
QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
Db 1287 GCTTTCTGTCACCAAGAGTTCAGCAGCTCATGGAGTCTCTGAGGCTCTATGGAGCC 1346
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1347 TTTCTGGCTTGCCTCCGCTTTCGATCTTCAGTGTCTTTTACATTTTCTGTGTCT 1406
QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1407 GAACTAAAGGAAGACTCTGGAACAAATCAGAGCCCATTTTGGAGGGCGA 1457

RESULT 6
AY208940
LOCUS
DEFINITION Bos taurus glucose transporter 8 (GLUT8) mRNA, complete cds.
ACCESSION AY208940
VERSION AY208940.1 GI:37784545
KEYWORDS
SOURCE Bos taurus (cow)
ORGANISM Bos taurus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae;
Bovinae; Bos.
REFERENCE 1 (bases 1 to 2073)
Zhang, F.-Q., Dong, B. and Zheng, Y.
Molecular cloning and expression of GLUT8 in bovine tissues
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 2073)
Zhang, F.-Q.
Direct Submission
JOURNAL
Submitted (31-DEC-2002) Department of Animal Science, University of
Vermont, 219 Terrill, 570 Main Street, Burlington, VT 05405, USA
FEATURES
Location/Qualifiers
1..2073
/organism="Bos taurus"
/mol_type="mRNA"
/db_xref="taxon:9913"
/tissue_type="kidney"
1..2073
/genes="GLUT8"
96..1532
/genes="GLUT8"
/notes="bGLUT8"
/codon_start=1
```

```
/product="glucose transporter 8"
/protein_id="AAP43920.1"
/db_xref="GI:37784546"
/tranlation="MTPEDQETOPLLRPPGGSAPRGREVFLEAAFAAALGPLSGFAL
GYSSPAISLRRAAPPAHLDEDAASWFGAIYVTLGNAAGSVLGGHLLDRAGKLSLVL
CALPFAVAVITAAQNLMLLGGRLLTGLACIASLVAPYVISEIAYPEVGLLQSC
VQMVVTGILLAYLAGWLEWLAVLCVPPSFMILLMFCMPETPRFLLSQHKHQA
NAAMQFLMGYACQWEEPLGAQDFHVAQLRRPGVYKFFIIGISLMAFQQLSGVNAV
FASGTIFEEAKFDKSSLASVVGVIQVLTATALLIMDRAGRLILLTILSGVVMVFT
SAFYVFKLTGGGPNSSHVDLPALVSMEDATNGLAWLAVGSMCLFTAGRAVGWGP
IPWLLMSIFPLHVGATGCVLTNFWFALVTKRFSLSMEVLRFYGGAFWLASAFCI
FGVLFTLACVPETPKTKLEQITAHFGR"
ORIGIN
Alignment Scores: 1..3e-150 Length: 2073
Pred. No.: 2230..50 Matches: 425
Percent Similarity: 94.35% Conservative: 26
Best Local Similarity: 88.91% Mismatches: 26
Query Match: 90.78% Indels: 1
DB: 4 Gaps: 1
US-09-886-954A-1 (1-477) x AY208940 (1-2073)
QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 96 ATGACGCTGAGGACCAAGAGGAGACCCAGCGCTCTTACGCGCGCGCGCGCGCT 155
QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
Db 156 CCCCAGCGCGCGCGCGCTTCTCTGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCT 215
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db 216 GGCCTCGCCCTCGCTACAGCTCCCGCGCCATCCCGAGCTTGGAGCGCGCGCGCGCG 275
QY 61 AlaProArgLeuAspAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAlaAla 80
Db 276 GCGCGCGCGCGCGCGCGCTCTCTGCTTTCGCGCGCGCGCGCGCGCTCTGCGCGCGCG 335
QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db 336 GCGCGCGCGCGCGCGCTCTGCGCGCTTCTGCGCGCGCGCGCGCGCGCGCGCGCGCT 395
QY 101 LeuLeuCysSerValProPheValAlaAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 396 GTCTCTGCGCGCTTCTGCGCGCTTCTGCGCGCTTCTGCGCGCTTCTGCGCGCTTCTGCG 455
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 456 TGGATGCTGCTCGGAGCGCGCTGCTCACCGCGCTGGCGCTGGCGCGCTTCTGCTGCTG 515
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgLysLeuGlySerCys 160
Db 516 GCGCGCGCTATATCTCTGAAATTCCTACCTGAGGTGCGAGGCTGCTGCGCTCTCTGT 575
QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db 576 GTGAGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 635
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 636 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 695
QY 201 PheMetProGluThrProArgPheLeuThrGlnHisArgArgGlnGlnAlaMetAla 220
Db 696 TTCAATGCTGAGACCCCTGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 755
QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
Db 756 GCAATGAGTCTCTGCTGGGCTTATGCCAGGGTGTGGAGAGCTTCTCTCTCTCTCTCTCT 815
QY 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleTyrIleTyrIle 259
```

/db xref="GI:12711201"	
/translation="MSPEDPTQPTPLRSPGARAPGRRVFLATFAAALGPLSLFSGFAL	
GYSFPAISRRITAPALRLGDTAASWFGAVVTGLGAAGVGLGMLLDRAGRKLSLL	
CTVPTGFAVITTAARDVWMLLGGRLTGLACGVASLVAPVYISEIAYPAVRGLGSC	
VQMLVVTGLLAVVAGVWRLVGLCVPTLMLLMCIMPETPRFLITQHOYQEA	
VQMLRLWSEEGEPPVGAHQFQGLAMLRPGVHKPLIIGICLMVFOQLUSGVNAI	
MFVANTIPEAKFSDSLASVTGIIQVLFPTAVAHLMIDRAVGLKLLALSGLVTFPSM	
SARFTYKLTQSGPSNHHVGLVPIASAPADVHLGLAWLAVGSMCLFTAGFAVWGFP	
I PWLLMSEIFPLHIKGVATGVCVLTNNWNAFLVTKEFNSIMEILRPYGAFWLTAAFCI	
LSVLFILITLTFVPTKTRLEQITAHFEGR"	
ORIGIN	
Alignment Scores:	
Pred. No.:	9,86e-146
Score:	2162.50
Percent Similarity:	92.26%
Best Local Similarity:	85.36%
Query Match:	88.01%
DB:	6
Length: 2087	
Matches: 408	
Conservative: 33	
Mismatches: 36	
Indels: 1	
Gaps: 1	
US-09-886-954A-1 (1-477) x AX076669 (1-2087)	
Qy	1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db	30 ATGTGCGCCGAGGACCCCAAGAAACACAGCCGCTGTTCGGTCCGCGGCGCCAGGGCT 89
Qy	21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db	90 CCGCGGCGCCGCGGCTTCCTTCCACCTTTCGCGCGCCGCTGGGCGCCGCTCAGCTTC 149
Qy	41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db	150 GGCCTTCGCGCTTGGCTACAGCTCCCGCCCATCCAGCCCTGCGGCGCCAGCCGCCCT 209
Qy	61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db	210 GCCCTAGCCCTCGAGACACTCGCGCCCTCTCGTTCGCGGCGCTGTCGACCTGGGCGCT 269
Qy	81 AlaAlaGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
Db	270 GCGCAGGGGCGTGGTGGCGGCTGCTGCTGACCTGCGGCGCCAGCTGAGGCTC 329
Qy	101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db	330 CTGCTCTGACCCGCTTCGTCACCGCTTTCGTCATCACCGCGCCCGCGGAGATGTG 389
Qy	121 TrpMetLeuLeuGlyGlyValArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db	330 TGGATGCTGCTCGAGGCGCGCTCTCTCACCGCTAGCTCGCGGAGTGCCTCATTTAGTG 449
Qy	141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db	450 GCACCGGTCTATATCTCGAAATCGCTTACCCAGCCGTTCCAGGAGTGTCTCGGCTCTGT 509
Qy	161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db	510 GTGACGCTGATGTGTCTGCGGCACTCTTGGCTACGTCGAGGCTGGTCTCTAGAA 569
Qy	181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db	570 TGGCGCTGTTGGCTGTCTGGGCTGTGTGCGCCCGCCACCTCATGCTCTGCTCATGTGC 629
Qy	201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGlnAlaMetAla 220
Db	630 TACATGCCCGAGACCCACGCTTCTCTGACGCAACACACAGTACACGAGGCGCATGCT 689
Qy	221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
Db	690 GCGCTTGGCTTCTGTGGGCTCTGAGGAGGCTGGGAGAGCCCTCTGTTGGGCTGAG 749
Qy	241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLeuProPheIleIle 259
Db	750 CACGAGGGCTTCCAGCTGGCTATGTGTGAGGCGCCCTGTGTGTCCACAGCCCTCATCATC 809


```
Db 446 GCACCGGTTTACATCTCGGAATTCGCTACCCAGCTGTCCGAGGACTGTCTCGCTCTCTGT 505
Qy 161 ValGlnLeuMetValValValGlyValLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db 506 GTGACAGTGTATGTTGCTACGTGCATCTCTCTGCTGTATGTGCGAGCTGGTCTTAGAG 565
Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 566 TGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 625
Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db 626 TACATCCCGAAACCCACGCTTTCTCTCACTCAACACAGTACCAGGAGGCGCATGGCT 685
Qy 221 AlaLeuArgPheLeuTrpGlySerGlnGlnGlyTrpGluAspProProLeuGlyAlaGlu 240
Db 686 GCTTGGCTTCTGTTGGGCTCTGAGAGGGCTGGAGAGCCCTGTGTGGGCTGAG 745
Qy 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyValLeuPheLeuLeu 259
Db 746 CACCAGGGCTTCCAGCTGGCCCTGCTGAGGCGCTTGGCATCTACAAGCCCTCATCATC 805
Qy 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
Db 806 GGCATTTCCTCATGCTTTCACAGAGCTGTCCAGGGGTCAATGCTCATGTTCTATGCC 865
Qy 280 GluThrIlePheGluGluAlaValPheLeuAspSerSerLeuAlaSerValValValGly 299
Db 866 AACAGCATCTTCAGAGAGGCAAGTTCAAGAGCAGCAGCTGGCCCTCGGTCACTGTGGGC 925
Qy 300 ValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArg 319
Db 926 ATAATCCAGCTCTGTTCACTGCTGTGGGCGCTCATCATGACAGAGAGGCGGAGG 985
Qy 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
Db 986 CTGCTCTGCGCTTGTGCGGTGTGATCATGCTGATGATGATGATGATGATGATGATGAT 1045
Qy 340 PheLeuLeuThrGlnGlyGlyProGlyAenSerSerHisValAlaIleSerAlaProVal 359
Db 1046 TTCAAATGACCCAGAGCTCCCAAGCACTCTCCCAAGTGGCTG---GTGCCCATC 1102
Qy 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
Db 1103 GCGCGGAGCTGTGATGTCCAGTGGGACTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1162
Qy 380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProIleProIleLeuMetSerGlu 399
Db 1163 CTCTTCATTGCTGGCTTGTGGGTGGCTGGGAGCCATCCCTCGGCTCTCATGTCTAGAG 1222
Qy 400 IlePheProLeuHisValValGlyValAlaThrGlyValCysValLeuThrAsnTrpLeu 419
Db 1223 ATCTTCCCTGTCATGTCAAGGTGTGGGTACCGGCATCTGTGTCTCCACCACTGGTTC 1282
Qy 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
Db 1283 ATGGCTTCTTAGTGACCAAGAGTTCAGAGGCTCATGGAGATGTCTAGACCTACGCT 1342
Qy 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
Db 1343 GCCTTCTGGCTACCCGCTGCTTCTCGCTCTCAGTGTCTCATGCTATTTCACACTGACCGTGT 1402
Qy 460 ProGluThrLysGlyValThrLeuGlnIleThrAlaHisPheGluGlyArg 477
Db 1403 CTTGAGACTAAAGGCGAGGACTCTGGAAACAAGTCAACAGCCCATTTTCAGGGAGCA 1456

RESULT 10
AF232061
LOCUS Mus musculus glucose transporter GLUT8 mRNA, linear ROD 14-JUL-2004
DEFINITION AF232061
ACCESSION AF232061
VERSION AF232061.2 GI:50295426
```

```
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1795)
AUTHORS Carayannopoulos,M.O., Chi,M.M., Cui,Y., Pingsterhaus,J.M.,
McKnight,R.A., Mueckler,M., Devaskar,S.U. and Moley,K.H.
TITLE GLUT8 is a glucose transporter responsible for insulin-stimulated
glucose uptake in the blastocyst
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 97 (13), 7313-7318 (2000)
MEDLINE 20319023
PUBMED 10860996
REFERENCE 2 (bases 1 to 1795)
AUTHORS Moley,K.H., Carayannopoulos,M.O. and Cui,Y.
TITLE Direct Submission
JOURNAL Submitted (07-FEB-2000) OB/Gyn, Washington University, 4566 Scott
Ave, St. Louis, MO 63110, USA
COMMENT On Jul 14, 2004 this sequence version replaced gi:8671757.
FEATURES
source
1..1795
/organism="Mus musculus"
/mol_type="mRNA"
/strain="129"
/db_xref="taxon:10090"
1..1434
/codon_start=1
/product="glucose transporter GLUT8"
/protein_id="AAF78366.1"
/db_xref="GI:8671758"
/translation="NSPEDPQETPLRPPEARTPRGRVFLASFAAALGPLSPGFAL
GYSPALPSLRRTAPPALRDGNAASWFGAVTTLGAAGGILGMLDPSGRKLULL
CTVPVGTGAVITTAARDVWMLGRLTGLACGVALVAPVVISETAYPAVRGLGSC
VQLAVTGLLAVAGVLEWRWLVLCVPTLMLLNLCYMPETPRFLITQHQYQEA
MAALPLMGSEGEWPEPVGAHQGFQALRRPGYIKPLIIGISLWVQQQSGVNAI
MFYANSIFEEKAFKDSLSAVTVGIQVLFATAALIMDRAGRLLLALSGVIMVFSM
SAFYTFKLQSLPSNSHVLVPIAAEPDVQVGLAWLAVGSMCLFIAGFAVGMGP1
PWLMSSETPLHLVKGATGICVLTWNFNAFLTKFSSVWEMLRPYGAFWLTAAPCAL
SVLETLTVVPTKGRITLSEQVTAHFEGR"
1435..1795
3'UTR
ORIGIN
Alignment Scores:
Pred. No.: 2,41e-145 Length: 1795
Score: 2156.00 Matches: 411
Percent Similarity: 92.47% Conservatives: 31
Best Local Similarity: 85.98% Mismatches: 34
Query Match: 87.75% Indels: 2
DB: 10 Gaps: 2
US-09-886-954A-1 (1-477) x AF232061 (1-1795)
Qy 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 1 ATGAGTCCGAGGACCCCGAGAGACTCAGCGCTATTTCGGCGCCACCGGAAGCCAGACT 60
Qy 21 ProArgGlyValArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
Db 61 CCCCCTGGCGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 120
Qy 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db 121 GCTTTCGGCTCGGCTACAGCTCCCGCCATCCCGAGCTCGGCGGCGCACCCGCCCCG 180
Qy 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db 181 GCCTCTGGCGCTCGAGACAATCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 240
Qy 81 AlaAlaGlyGlyValLeuGlyValTrpLeuValAspArgAlaGlyArgGlySerLeu 100
Db 241 GCTCAGGGGGCATACTCTGGCGGCTGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 300
Qy 101 LeuLeuCysSerValProPheValAlaAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
```

```

Db 1378 CTGAGACTAAAGCAGGACTCTGGAAACAGTACAGCCCAATTCGAGGAGCA 1431
RESULT 11
LOCUS AX076671 2072 bp DNA linear PAT 06-FEB-2001
DEFINITION Sequence 5 from Patent WO0104145.
ACCESSION AX076671
VERSION AX076671.1 GI:12711202
KEYWORDS
SOURCE Mus sp.
ORGANISM Mus sp.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1
AUTHORS Thorens,B., Ibberson,M. and Uldry,M.
TITLE Glutx polypeptide family and nucleic acids encoding same
JOURNAL Patent: WO 0104145-A 5 18-JAN-2001;
University of Lausanne (CH)
FEATURES
source Location/Qualifiers
1..2072
/organism="Mus sp."
/mol_type="unassigned DNA"
/db_xref="taxon:10095"
21..1454
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC28497.1"
/db_xref="GI:12711203"
/translation="MSPEDQETQPIRRPEARTPRRRVFLASFAAALGLNFGPAL
GYSPALPSRRRTAPALRGDNDASWFGAVTLGAAAGGILGGWLLDRSGKLSLL
CTVPFVTFVAVITAAARDVWMLLGGRLTLGACVAVSLVAVPISEIAYPAVRGLLQSC
VOLMVTGIIILAVAGVLEWRWLAVLVCVPTMLLMCMYPTPRFLUTHQYQEA
MAALPLMGSEGEWEPVGAHQFOLLRRPGYKPLIIGISLVFQOLSGVNAI
MFYANSIPEAKKSSLSAVTGIIQVLFVAALIMDRAGRLILALSGVIMVFSM
SFATYFKLQSLPSNSSHVGLPIAAEPVDVQGLAWLAVGSMCLFIAGFVWGSP
PWLMSSEIFPLHVRGVAIGICVLINFMFALVTKFNSVMEMLRPYGAPWLTAFCAL
SVLFTLTVPETKRTLEQVTAHFEH"
ORIGIN
Alignment Scores: 7.69e-145 Length: 2072
Pred. No.: 2150.00 Matches: 409
Score: 92.47% Conservat: 33
Percent Similarity: 85.56% Mismatches: 34
Best Local Similarity: 87.51% Indels: 2
Query Match: 6 Gaps: 2
DB:
US-09-886-954A-1 (1-477) x AX076671 (1-2072)
QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 21 ATGTCTCCGAGGACCCCGAGGAGCAGCGCGCTATTTCGCGGCCACCGAAGCCAGGACT 80
QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db 81 CCCCAGCGCGCGCGGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 140
QY 41 GlyPheAlaLeuGlyTyxSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db 141 GCCTTCGGCTCGGCTACAGCTCCCGCGCATCCCGCGCTCCCGCGCGCGCGCGCGCG 200
QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db 201 GCCTCGCGCTCGGAGACAATCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 260
QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
Db 261 GTTCAGCGGGGATACTCTGCGCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 320
QY 101 LeuLeuGlySerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 321 TTGCTCTGACCGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 380

```

```

QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 381 TCGATGCTGCTGGAGGCGCTCTCTCACGGGCTTGGCTGGAGTCGCTCCTCCTCCTAGTG 440
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db 441 GCACCGGTTTACATCTCGGAATCGCTTACCCAGCTGTCCGAGGACTGCTCGGCTCCTGT 500
QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu 180
Db 501 GTGCAGCTGATGGTTGTCTACTGGCATCTCTGGCTATGTGGCAGGCTGGTCTCCTAGAG 560
QY 181 TrpArgTyrLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 561 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 620
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
Db 621 TACATGCCCGAGACCCAGCTTTTCTCTCACTCAACACCACTGATACAGGAGCCATGGCT 680
QY 221 AlaLeuArgPheLeuTyrGlySerGluGlnGlyTyrGluAspProProIleGlyAlaGlu 240
Db 681 GCCTTGGCTTCTGTGGGCTCTGAGGAGGCTGGGAGAGCCCTGTGTGGGCTGAG 740
QY 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleProPheIleIle 259
Db 741 CACGAGGCTTCCAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 800
QY 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
Db 801 GGCATTTCCTCATGCTTCCAGCAGCTGTGAGGCTCAATGCTATCATGTTCTATGCC 860
QY 280 GluThrIlePheGluGluAlaValPheLysAspSerSerLeuAlaSerValValValGly 299
Db 861 AACAGCTTCTCAGAGGCGCAAGTTCAAGAGCAGCTGGCTGGCTGGCTGGCTGGCTGGCT 920
QY 300 ValIleGlnValLeuPheThrAlaValAlaLeuLeuMetAspArgAlaGlyArgArg 319
Db 921 ATAATCCAGCTCTGTTCACTGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 980
QY 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
Db 981 CTGCTCTCGGCTTGTGGGCTGTGATCATGGTGTTCAGTATGAGTGGCTTGGTACCTAC 1040
QY 340 PheLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProVal 359
Db 1041 TTCAAACTGACCCAGAGCTCCCACTCTCCAGTGGCTG---GTGCCCATC 1097
QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCys 379
Db 1098 GCGGCGAGCTGTGATGTCAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1157
QY 380 LeuPheIleAlaGlyPheAlaValGlyTyrGlyProIleProIleProIleLeuMetSerGlu 399
Db 1158 CTCTTCACTGCTGGCTTGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1217
QY 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrIleThrLeu 419
Db 1218 ATCTTCTCTGTCATGTCAGGCTGTGGCTGTACGGCTATGTCCTCAACCACTGGTTC 1277
QY 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
Db 1278 ATGGCTTTCTAGTGACCAAGGTTCAACAGGCTCATGGAGTGTCTGAGACCTACGCT 1337
QY 440 AlaPheTyrLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
Db 1338 GCCTTCTGCTACCCGCTGCTTCTGGCTCTCAGTGTCTTATTCACATGACCGTGTGC 1397
QY 460 ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1398 CTGTGACTAAAGGCGAGGCTCTGGAAACAAGTCAACAGCCCATTTTCAGGGGACGA 1451

```

RESULT 12

```

MMU245936
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS
TITLE
JOURNAL

```

```

FEATURES
source

```

gene

CDS

```

MMU245936      2072 bp      mRNA      linear      ROD 18-FEB-2000
Mus musculus mRNA for glucose transporter (GLUTX1 gene).
AJ245936
AJ245936.1     GI:7018606
Mus musculus glucose transporter; GLUTX1 gene.
Mus musculus (house mouse)
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Ibberson, M., Uldry, M. and Thorens, B.
GLUTX1, a novel mammalian glucose transporter expressed in the
central nervous system and insulin-sensitive tissues
J. Biol. Chem. 275 (7), 4607-4612 (2000)
20138191
10671487
2 (bases 1 to 2072)
Ibberson, M.R.
Direct Submission
Submitted (02-SEP-1999) Ibberson M.R., Institute of Pharmacology
and Toxicology, University of Lausanne, Rue du Bugnon 27, Lausanne,
CH-1005, SWITZERLAND
Location/Qualifiers
1..2072
/organism="Mus musculus"
/mol_type="mRNA"
/db_xref="taxon:10090"
1..2072
/gene="GLUTX1"
21..1454
/gene="GLUTX1"
/codon_start=1
/product="glucose transporter"
/protein_id="CAB75719.1"
/db_xref="GI:7018607"
/db_xref="GOA:Q9JIF3"
/translation="NSPDPQPTQPLLRPPPEARTRPRGRVFLASPAALGPLNFGPAL
CYSSPAISLRTPAPPALRLGDNAASWFGAVVTLGAAAGGILGGWLLDRSGKLSLLL
VOLVVTVGILLYAVAGVLEWRVLAVLCVPTMLLMCMYPTPRFLTLTHQYQEA
MNAALFLGSEEGWEEPPGAGHQFQALLRPGYKPLIIGISLMVQQSLSGVNAI
MFYANSIFEEAKFKDPSLASVTGIIQVLFATAVALIMDRAGRRLLALSGVIMVFSM
SARGYFKLTQSLPSNSHVGVPVIAEPVDVQVGLAVGSMCLFIAGFAVGWGPFI
PWLMSLFFPLHVKGVATGICVLITWFWAFLVTKFNSVWEMLRPYGAFWLTAAFCAL
SVLTLTIVVPTKGTLEQVTAHFGR"

```

ORIGIN

```

Alignment Scores:
Pred. No.:      7,69e-145      Length:      2072
Score:          2150.00      Matches:      409
Percent Similarity: 92.4%      Conservative: 33
Best Local Similarity: 85.5%      Mismatches: 34
Query Match:      87.51%      Indels:      2
DB:              10          Gaps:      2

```

US-09-886-954A-1 (1-477) x MMU245936 (1-2072)

```

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
Db 21 ATGTCTCCGAGACCCCGAGGACGCGCTATTTCGGCCCGGAGCCAGGACT 80
QY 21 ProArgGlyArgArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db 81 CCCCCTGGCGCGCGGCTCTTCTCGCTTCTTCGCGCGCGCTCTGGGACCTCACTTC 140
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnAlaAlaProPro 60
Db 141 GCGTTCGGCTCGGCTACAGCTCCCGCCCATCCCGAGCTCGGCGGACCGGACCCCG 200
QY 61 AlaProArgLeuAspAlaAlaAlaSerTyrPheGlyValValThrLeuGlyAla 80

```

Db 201 GCCCTGGCGCTCGAGACAATCGCGCTCTCTGTTTCGGGGCGCTCGTGACCCCTGGCGCT 260
Qy 81 AlalaGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db 261 GCTGAGGGGATATAGTGGCGGCTGGCTCTGGAGCCGTTGAGGCGCAAGCTGAGCCTC 320
Qy 101 LeuLeuGlySerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 321 TTGCTCTGGACCGTGGCTTCGTGACTGGCTTGTGTCTCATCCCGCGCGCGGATGTG 380
Qy 121 TrpMetLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 381 TGGATGCTGCTCGAGGCGGCTCTCTACCGGCTTGGCTGGAGTGGCTCACTAGTG 440
Qy 141 AlaProValTyrIleSerGluLeuLeuAlaTyrProAlaValArgGlyLeuGlySerCys 160
Db 441 GCACCGGTTTACATCTCGGAATCGCTTACCAGCTGTCCGAGGACTGCTCGGCTCTGT 500
Qy 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db 501 GTGAGCTGATGTTGTCTAGTGGCTCTCTGGCTATGTGGAGGCTGGGTCTCTAGAG 560
Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 561 TGGCGCTGGCTGGCGCTGGCTGTGTGGCCGCCACCCTCATGCTGCTGCTCATGTGC 620
Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAla 220
Db 621 TACATGCCCGAGACCCCGCTTTCTCTCACTCAACACCATGATACAGGAGCCATGGCT 680
Qy 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu 240
Db 681 GCTTGGCTTCTGTGGGCTCTGAGGAGGCTGGAGAGCCCTCTGTGGGCTGAG 740
Qy 241 --GlnSerPheHisLeuAlaLeuArgGlnProGlyIleTyrIleProPheIlele 259
Db 741 CACAGGGCTTCAGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 800
Qy 260 GlyValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
Db 801 GGCATTTCCTCATGCTCTTCAGCAGCTGTGAGGGTCAATGCTATCATGTTCTATGCC 860
Qy 280 GluThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValGly 299
Db 861 AACAGCATCTTCAGAGAGCCAGTTCAGAGCAGCAGCTGGCTGGCTGGCTGGCTGGCT 920
Qy 300 ValIleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArg 319
Db 921 ATAATCCAGGCTCTGTCTACTGCTGGCGGCTCTCATGAGCAGCTGGCTGGCTGGCTGGCT 980
Qy 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
Db 981 CTGCTCTGGCTTGTGGGTTGTATCATGGTGTTCAGTATGAGTGGCTTGGTACCTAC 1040
Qy 340 PheLysLeuThrGlnGlyProGlyAenSerSerHisValAlaIleSerAlaProVal 359
Db 1041 TTCAACTAGCCAGGAGCTCCCGCACTCTCCAGTGGCTGGCTGGCTGGCTGGCTGGCT 1097
Qy 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
Db 1098 GGGGCGGAGCTGTGGATGTCAGAGTGGAGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1157
Qy 380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProIleProIleLeuMetSerGlu 399
Db 1158 CTCCTTCATGCTGGCTTGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 1217
Qy 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
Db 1218 ATCTTCTCTGATGTCAGGGTGTGGCTACCGGATCTGTCTCTCACTCACTGGTTC 1277
Qy 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
Db 1278 ATGGCTTCTTAGTGACCAAGAGTTCAACAGCGTCTATGGAGTGTCTAGACCTACGGT 1337

Qy 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
Db 1338 CCTTCTGGCTACCGCTGCCCTTCTGGCTCTCAGTCTCTATTCACACTGACCGTTGTC 1397
Qy 460 ProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1398 CCTGAGACTAAAGCAGGACTCTGGAACAAGTCAAGCCATTCGAGGAGCA 1451

RESULT 13

AB033418
LOCUS
DEFINITION Rattus norvegicus glut8 mRNA for glucose transporter 8, complete cds.
2189 bp mRNA linear ROD 11-APR-2000

AB033418
ACCESSION
VERSION AB033418.1 GI:7592743
KEYWORDS glucose transporter 8.
SOURCE Rattus norvegicus (Norway rat)
ORGANISM Rattus norvegicus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae;

REFERENCE
1 (bases 1 to 2189)
AUTHORS Ishibashi,K.
TITLE Molecular cloning of a new putative glucose transporter
JOURNAL Published Only in DataBase (2000)
REFERENCE
2 (bases 1 to 2189)
AUTHORS Ishibashi,K.
TITLE Direct Submission
JOURNAL Submitted (08-OCT-1999) Kenichi Ishibashi, Jichi Medical School,
Pharmacology; Minami-kawachi, Kawachi, Tochigi 329-0498, Japan
(E-mail:kishiba@jichi.ac.jp, Tel:81-285-58-7326,
Fax:81-285-44-5541)

FEATURES

Location/Qualifiers
1..2189
/organism="Rattus norvegicus"
/mol_type="mRNA"
/db_xref="taxon:10116"
/clone_lib="testis"
1..2189
/gene="glut8"
/note="Glut8 has 12 transmembrane segments"
/codon_start=1
/product="glucose transporter 8"
/protein_id="BA094383.1"
/db_xref="GI:7592744"
/translation="MSPEDPQETQPLRSFGARAPGRRVFLATFAALGPLSGFAL
GYSPALPSLRRTAPPALRLDGTAAWFGAVVTLGAAAGQAVLGLLDRAGRLSL
LCTVFTGPAVITAARDVWMLLGGRLITGLACGVALVAPVYISEIAYPAVRGLGS
CVQLMVTGILLAYVAGVLEWRMLVLCVPTLMLLMCMYMPETPFLTHOYQOE
AMALRFUNGSEGEPEFVGAHQGQFQLAMRLRPVGHKPLIIGICLMVFQQLSGVNA
IMFYANTIFEAKFKDSLSAVTGGIIQVLFATAALIMDRAGRLKLLALSGVINVFS
MSAGTPTKQSGPSNSHGLLVPIISAEPADVHLGLAWLAVGSMCLFIAGFVGWG
PIPLLMSKIEPLIKGVATGCVLTWNFMFLVTKFNSIMEILRPYGAFLTAFC
ILSLVLTITFTVPETKGRLEQITAHFGR"

ORIGIN

Alignment Scores:
Pred. No.: 1.59e-144 Length: 2189
Score: 2146.00 Matches: 407
Percent Similarity: 91.86% Conservative: 33
Best Local Similarity: 84.97% Mismatches: 37
Query Match: 87.34% Indels: 2
DB: 10 Gaps: 2
US-09-886-954A-1 (1-477) x AB033418 (1-2189)

Qy 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 109 ATGTGCCCCGAGGACCCCGAGAAACACAGCCGCTGTTGCGGTCCCGGCGCCAGGCT 168


```
QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
|||
DB 169 CCGCGGCGCGCGGGTCTTCCTTGGCCACCTTCGCGCGCGCCCTGGGCGCGCTCAGCTTC 228
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
|||
DB 229 GCGTTCGCGGCTTGCTACAGCTCCCGCGCATCCCGAGCCTGCGGCGCACCGCACCCCT 288
QY 61 AlaProArgLeuAspAlaAlaAlaSerTyrPheGlyValAlaValValThrLeuGlyAla 80
|||
DB 289 GCCTACGCTCGGAGACACTGGGCGCTCTGGTTCGGGGCGCTCGTACCTTGGGCGGT 348
QY 81 AlaAla---GlyGlyValLeuGlyGlyTyrLeuValAspArgAlaGlyArgTysLeuSer 99
|||
DB 349 GCGGCGCGAGGGGCGCGTCTGGCGCGCTGCTCTCGACCGCTCGAGCGCGCAAGCTGAGC 408
QY 100 LeuLeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAsp 119
|||
DB 409 CTCTGCTCTGACCGTGGCCCTTCGTGACCGGTTTTGCTGTCATCCCGCGCGCGGAT 468
QY 120 ValTrpMetLeuLeuGlyGlyArgLeuLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeu 139
|||
DB 469 GTGTGATGCTGCTCGAGGCGCGCTCTCTCACCGGCTAGCTCGGAGTTCGCTCATTA 528
QY 140 ValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuGlySer 159
|||
DB 529 GTGGCACCGGCTCATATCTCGGAATCGCTACCCAGCGGTTCCAGGACTGCTCGGCTCC 588
QY 160 CysValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeu 179
|||
DB 589 TGTGTGAGTGAATGTTGCTACGTGGCATCTCTTGGCCCTACGTGGGCGGTGGGCTCA 648
QY 180 GluTrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMet 199
|||
DB 649 GAATGGCGTGTGTGGCTGTGGCTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGG 708
QY 200 CysPheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMet 219
|||
DB 709 TGTACATGCCGAGACCCCGCTTCTCTCTGACGCAACACCGTACCAGGAGGCGCATG 768
QY 220 AlaAlaLeuArgPheLeuTrpGlySerGluGlnGlyTyrGluAspProIleGlyAla 239
|||
DB 769 GCTGCGCTCGCTCTGTTGGGCTCTGAGGAGGCGTGGGAGAGAGAGAGAGAGAGAGAG 828
QY 240 Glu---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleValProPheIle 258
|||
DB 829 GAGCACAGGGCTTCAGCTGGTATGCTGAGGCGGCTGTGGTGTCCACAGGCGCTCATC 888
QY 259 IleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyr 278
|||
DB 889 ATCGCATTTGCTCATGCTTCTCCAGCAGCTGTCCAGGGGTCAACGCCATCATGTTCTAT 948
QY 279 AlaGluThrIlePheGluGluAlaIlePheLysAspSerSerLeuAlaSerValValVal 298
|||
DB 949 GCCAACACCATCTTTTCAGAGGCGCAATTCAGAGCAGCAGCAGCAGCAGCAGCAGCAG 1008
QY 299 GlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArg 318
|||
DB 1009 GGCATCATCCAGGTTCTGTTCTGCTGTGGGCGGCTCATCATGTGACAGAGCAGGCGCA 1068
QY 319 ArgLeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAla 338
|||
DB 1069 AAACCTGCTTCGGCTTGTGGGTGTGATCATGGTTTTTCAGCATGAGCGCTTTGTGTACC 1128
QY 339 TyrPheIleValThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaPro 358
|||
DB 1129 TACTTCAAACTGACCGAGTGGCCCGCAACTCTCTCCCATGTAGCGCTCTCTGGTGCGCC 1188
QY 359 ValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMet 378
|||
DB 1189 ATCTCCGAGAGCGCTGCTGATGTTTCACCTGGGGCTGGCTGTGGTGTAGGAGCAATG 1248
QY 379 CysLeuPheIleAlaGlyPheAlaValGlyTyrProIleProTrpLeuLeuMetSer 398
```

```
DB 1249 TGCTCTTCATCGCTGGTTTTGTCAGTAGGCTGGGAGCCATCCCTCGCTCATGTCA 1308
QY 399 GluIlePheProLeuHisValIleValAlaThrGlyIleCysValLeuThrAsnTyr 418
|||
DB 1309 GAGATCTTCCCTCTGTCACATCAAGGCTGTGGCTACCGGCTGTGTCTCTCCACCACTGG 1368
QY 419 LeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyr 438
|||
DB 1369 TTTCATGGCCTTCTGGTGACCAAGAGTTTACAGCATCATGGAGATCTTCAGACCTTAC 1428
QY 439 GlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCys 458
|||
DB 1429 GCGGCGCTTCTGGCTCACCCTGCTCTGTATCTCTCAGCGCTCTTTTCACGCTCACCCTT 1488
QY 459 ValProGluThrGlyGlySerThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
|||
DB 1489 GTCCCTGAGACTAAAGGCGAGGACTCTGGAAACAATCACAGCCCATTTTCGAGGAGCG 1545

RESULT 14
AX191507
LOCUS AX191507
DEFINITION Sequence 29 from Patent WO0149728.
ACCESSION AX191507
VERSION AX191507.1 GI:15209697
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Kato, S. and Kimura, T.
TITLE Human proteins having hydrophobic domains and dnas encoding these
JOURNAL Proteins
Patent: WO 0149728-A 29 12-JUL-2001;
Protegene Inc. (JP); SAGAMI CHEMICAL RESEARCH CENTER (JP)
FEATURES
source
1. .1461
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
61. .849
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC51153.1"
/db_xref="GI:15209698"
/translation="MTPSDPEETQPLGPPGSGAPRRRVFLAFAALGLPLSGFAL
GYSPAIPSLQRAAPPAPRLDDAAASWFGAVVTLAGAAGVGLGGWLVDRAGKLSLL
CSVPFVAGFAVITAAQDYVMLLGGRLTGLACVASLVAPVYISIIAYAVRGLGSC
VOLMVVVGILLAYLAGVLEWRWLAVLCVPPSLMLLMCFMPTPRFLLTQHRROEA
APGLVRCGHGVQHECLRLRLQADPGWPNQLLARGHLGACLTCTAC"
ORIGIN
Alignment Scores:
Pred. No.: 2,496-102 Length: 1461
Score: 1554.50 Matches: 324
Percent Similarity: 67.8% Conservative: 1
Best Local Similarity: 67.6% Mismatches: 1
Query Match: 63.2% Indels: 153
DB: Gaps: 2
US-09-886-954A-1 (1-477) x AX191507 (1-1461)
QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
DB 61 ATGACGCGCGAGAGCCAGAGAAACCCAGCGCTCTTGGGGCGCTCTCGCGGACGCGG 120
QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
DB 121 CCGCGGCGCGCGGCTGCTTCTCTCGCGCTTCTCGCGCTTCTGGGCCCATCTCAGCTTC 180
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
```



```
Db 121 GCGTGTCTGGGCTGTGTGGCCCCCTCTCTTTCATGCTGTCTGTCTCATGTGTTTTCATGCCCTGAG 180
QY 205 ThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPhe 224
Db 181 AGCCCTCGCTTCTGTCTGTCTAGCAGACAGGAGGCCATGGCCGCGCATGCGAGTTTC 240
QY 225 LeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu---GlnSerPhe 243
Db 241 CTGTGGGGCTATGCCAGAGGTTGGAGAGGCTCCCTTGGGGCTCAGCACCAGGACTTC 300
QY 244 HisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGlyValSerLeu 263
Db 301 CAGCTGGCCAGCTGGGGGCTCTGTGTCTACAGGCCCTTCATCATCGGCATCTCCCTG 360
QY 264 MetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePhe 283
Db 361 ATGGCCCTTCAGCAGCTGTTCAGGCGTCAATGCCGTCTATGTTCTACGGGAGACCACTTT 420
QY 284 GluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyValIleGlnVal 303
Db 421 GAGGAGGCTAAGTTCAAGGACAGCAGCTGGCCCTCAGTCGTCTGTGGGCGTCATCCAGGTG 480
QY 304 LeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuVal 323
Db 481 CTGTTCACTGCCACCGCGCGCCCTGATCATGGACAGAGCTGGGCGAAGGCTCTCTTGACC 540
QY 324 LeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeuThr 343
Db 541 TTGTCAAGTGTGGTGTATGTGTTCAGCACAGGCGCTTTGGTACCTACTTCAAGCTGACC 600
QY 344 GlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaGlnPro 363
Db 601 GAGGGCGGCGCCAGCAACTCTCCACAGTGGACCTCCCGGGCGCTCTCTCCATGGAGGCC 660
QY 364 ValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAla 383
Db 661 GCGGATACCAACGTGGGGGCTGGCGTGGCTGGCGGTGGGCAACATGTGCGCTCTTCATCGCC 720
QY 384 GlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeu 403
Db 721 GGCCTTCGCTGTGGGCTGGGGGCCCATCCGTTGGCTCTCATGTACAGAGATCTTCCCTCTG 780
QY 404 HisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeu 423
Db 781 CATGTCAAGGGCGTGCCACCAGCGGCTGCGTCTCTCAACCACTGGTTCATGGCCTTTCTG 840
QY 424 ValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeu 443
Db 841 GTGACCAAAAGAGTTCAAGCAGTCTCATGGAGGTGCTCAGGCGCCCTACGGTGCCTTCTGGCTG 900
QY 444 AlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLys 463
Db 901 GCCTCTGGCTTCTGCAATCTGGTGTCTTTTTCACCTCTGGCGCTGTGTCTGAGACCAA 960
QY 464 GlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 961 GGGAGACTCTGGACAAATCACAGGCCATTTTGGAGGGCGGA 1002
```

Search completed: February 25, 2005, 20:31:23

Job time : 5130 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 25, 2005, 16:25:41 ; Search time 644 Seconds
(without alignments)
4384.653 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457

Sequence: 1 MTPEDPETOQLLPGPGGSA.....CVPETKGTLEQITAHFEGR 477

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlp
-O=/cn2.1/USPTD.spool.p/US09886954/runat.23022005.154241.13198/app_query.fasta_1.647
-DB=N_Geneseq.16Dec04 -OPT=fastap -SUFFIX=p2n.rng -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MAYRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALLIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09886954 @CN 1.1 708 @runat.23022005.154241.13198 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N_Geneseq.16Dec04.*

1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002as.*
7: Geneseqn2002bs.*
8: Geneseqn2003as.*
9: Geneseqn2003bs.*
10: Geneseqn2003cs.*
11: Geneseqn2003ds.*
12: Geneseqn2004as.*
13: Geneseqn2004bs.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2457	100.0	1445	8	Abz24794 Human sol
2	2457	100.0	1856	8	Abz24792 Human sol
3	2457	100.0	2080	4	Aad09552 Human tra
4	2448	99.6	1873	8	Abz24793 Human sol
5	2448	99.6	2217	4	Aaf55865 Human GLU

6	2162.5	88.0	2087	4	Aaf55866	Aaf55866 Rat GLUTX
7	2150	87.5	2072	4	Aaf55867	Aaf55867 Murine GL
8	2123.5	86.4	1954	6	Abx34029	Abx34029 Human can
9	2067.5	84.1	1923	13	ACN43514	ACN43514 Human dia
10	1912.5	77.8	1845	13	ACN43515	ACN43515 Human dia
11	1823.5	74.2	1790	11	ADL33401	ADL33401 Human tra
12	1746.5	71.1	1763	12	ADH45466	ADH45466 Human mol
13	1679.5	68.4	1156	6	ABL90126	ABL90126 Human pol
14	1554.5	63.3	1461	4	AAD12574	AAD12574 Human pro
15	972.5	39.6	2011	4	Aaf55871	Aaf55871 Rat GLUTX
16	958.5	39.0	2356	8	ADA84076	ADA84076 Human SLC
17	958.5	39.0	2356	8	ADA84076	ADA84076 Human SLC
18	958.5	39.0	2356	8	ADA84076	ADA84076 Human SLC
19	958.5	39.0	2356	8	ADA84076	ADA84076 Human SLC
20	958.5	39.0	2356	8	ADA84076	ADA84076 Human SLC
21	958.5	39.0	2356	8	ADA84076	ADA84076 Human SLC
22	958.5	39.0	2356	8	ADA84076	ADA84076 Human SLC
23	958.5	39.0	2356	8	ADA84076	ADA84076 Human SLC
24	958.5	39.0	2356	8	ADA84076	ADA84076 Human SLC
25	949	38.6	1541	4	Aaf55870	Aaf55870 Human GLU
26	904.5	36.8	2171	4	AAK94284	AAK94284 Human ful
27	904.5	36.8	2171	12	ADL30893	ADL30893 Full leng
28	853	34.7	1293	12	ADP28539	ADP28539 Human sec
29	819	33.3	1610	13	ACN41918	ACN41918 Human dia
30	804.5	32.7	1412	4	ABL13883	ABL13883 Drosophil
31	788.5	32.1	1724	4	ABL13389	ABL13389 Drosophil
32	788	32.1	2514	4	ABL17067	ABL17067 Drosophil
33	734	29.9	1662	4	ABL11183	ABL11183 Drosophil
34	705	28.7	1632	4	ABL02211	ABL02211 Drosophil
35	701.5	28.6	3716	4	ABL13882	ABL13882 Drosophil
36	687	28.0	9218	4	ABL17066	ABL17066 Drosophil
37	686.5	27.9	1527	2	AAZ11731	AAZ11731 Human tra
38	676	27.5	4471	3	ABC36319	ABC36319 Arabidops
39	670.5	27.3	1741	3	ABC36319	ABC36319 Arabidops
40	666.5	27.1	1614	4	ABL02481	ABL02481 Drosophil
41	666.5	27.1	1953	12	ADG47939	ADG47939 Corn Beta
42	666.5	27.1	3614	4	ABL02480	ABL02480 Drosophil
43	662.5	27.0	1683	3	ABC39099	ABC39099 Arabidops
44	661	26.9	4078	4	ABL11182	ABL11182 Drosophil
45	653.5	26.6	1755	3	AAC40459	AAC40459 Arabidops

ALIGNMENTS

RESULT 1
ABZ24794
ID ABZ24794 standard; cDNA; 1445 BP.
XX
AC ABZ24794;
XX
DT 07-APR-2003 (first entry)
XX
DE Human solute carrier type 2A nucleic acid.
XX
KW Human; solute carrier type 2A; SLC2A; glucose transporter; p53; cancer;
KW cytostatic; gene therapy; gene; ss.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT CDS 7..1440
FT /*tag= a
FT /product= "Human SCL2A"
XX
PN WO200298467-A1.
XX
PD 12-DEC-2002.
XX
PF 03-JUN-2002; 2002WO-US017419.
XX
PR 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.

XX (EXEL-) EXELIXIS INC.
XX Friedman L, Plowman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
XX WPI; 2003-201283/19.
XX P-PSDB; ABP58364.
XX Identifying candidate p53 pathway modulating agent as therapeutic target
XX for disorders of defective p53 function e.g. cancer, by assaying purified
XX solute carrier family 2 (SLC2A) polypeptide or nucleic acid with a test
XX agent.
XX Disclosure; Page 45; 58pp; English.
XX The present sequence is that of a human solute carrier type 2A (SLC2A)
XX nucleic acid. Genetic screens were designed to identify modifiers of the
XX p53 pathway in Drosophila in which p53 was overexpressed. Human
XX orthologues (polynucleotides and polypeptides) of one such modifier were
XX then identified, including the present nucleic acid. SLC2As are glucose
XX transporter proteins with sugar transporter domains. SLC2A nucleic acids
XX and polypeptides are attractive drug targets for the treatment of
XX pathologies associated with a defective p53 signalling pathway, such as
XX cancer. The invention provides in vitro and in vivo methods of assessing
XX SLC2A function. Modulation of an SLC2A or its binding partners is useful
XX for understanding the association of the p53 pathway and its members in
XX normal and disease conditions and for developing diagnostic and
XX therapeutic modalities of p53-related pathologies. SLC2A-modulating
XX agents that act by inhibiting or enhancing SLC2A expression, directly or
XX indirectly, e.g. by affecting an SLC2A function such as transport or
XX binding activity, can be identified using methods provided. Modulators
XX include small molecules, nucleic acids, antibodies, antisense
XX oligonucleotides and phosphothioate morpholino oligomers (claimed)
XX
SQ Sequence 1445 BP; 182 A; 500 C; 456 G; 307 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 1,328-204 Length: 1445
Score: 2457.00 Matches: 477
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 8 Gaps: 0

US-09-886-954A-1, (1-477) x AB224794 (1-1445)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
DB 7 ATGACGCCGAGGACCCAGAGGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGACGCG 66

QY 21 ProArgGlyArgValPheLeuAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
DB 67 CCCCAGCGCGCGCGCTCTCTCGCGCGCTTTCGCGCTTGGCGCTTGGCGCTTTCAGCTTC 126

QY 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
DB 127 GCGTTCGCGCTCGGTACAGCTCCCGCGCATCCCTAGCTCGAGCGCGCGCGCGCG 186

QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
DB 187 GCGCGCGCTGAGCAGCGCGCGCTCTCTGCTTGGGGCTTCTGAGACCTTGGGTGCC 246

QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlySerLeu 100
DB 247 GCGCGCGGGGAGTGTGGCGGCTGGCTGGTGGACCGCGCGCGGCGGCAAGCTGAGCCTC 306

QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
DB 307 TTGCTGTCTCCGTGCGCTTCTGGTGGCGGCTTTCGCTCATCACCGCGCGCGGACGCTG 366

QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
DB 367 TGGATGCTGTGGGGGCGCGCTCTCTCACCGCGCTGGCGCTGGCGGTGTTCCTCCCTAGTG 426

QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
DB 427 CCCCAGGCTATACATCTCGAAATCGCTTACCAGAGTCCCGGGGTGCTCGGCTCTCTGT 486

QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
DB 487 GTGAGCTAATGGTCTGCTGGCATCTCTCTCGCTTACCTGGCAGGCTGGGTGCTGAG 546

QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 547 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 606

QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAla 220
DB 607 TTTATGCCGAGAGCCCGCGCTTCTGCTGACTCAGCACAGCGCGCAGGAGCCATGGCC 666

QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
DB 667 GCGCTGGGTCTCTGTGGGGCTCCGAGCAGGGCTGGGAAGACCCCGCATCGGGCTGAG 726

QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleTyrProPheIleIleGly 260
DB 727 CAGAGCTTTACCTGGCGCTCTCGCGCAGCCCGGCATCTACAAGCCCTTCATCATCGGC 786

QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAlaAlaValMetPheTyrAlaGlu 280
DB 787 GTCTCCCTGATGGCTTCCAGCAGCTTCGGGGGTCAACGGCGTCACTGTTTATGACAG 846

QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
DB 847 ACCATCTTTGAAGAGGCCCAAGTCAAGGACAGCAGCTGGCTGGTGGTGGTGGTGG 906

QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuLeuMetAspArgAlaGlyArgArgLeu 320
DB 907 ATCCAGGTGCTGTTACAGCTGTGGCGCTCTCATATGAGCAGAGGCGGAGGCTG 966

QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
DB 967 CTCTCGGCTTGTGAGGTGGTCTATGGTGTTCAGCAGAGTGGCTTCGGCGCTACTTC 1026

QY 341 LysLeuThrGlnGlyProGlyLysSerSerHisValAlaIleSerAlaProValSer 360
DB 1027 AAGCTGACCCAGGGTGGCGCTGGCAACTCTCTCGCAGCTGGCCATCTCGCGCGCTGTCT 1086

QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
DB 1087 GCACAGCTGTGATGCGCAGCGTGGGGCTGGCGCTGGCTGGCTGGCTGGCTGGCTGG 1146

QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
DB 1147 TTATCGCGGCTTGTGGTGGCTGGGGGCCCATCCCTGGCTCTCATGTTCAGAGATC 1206

QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
DB 1207 TTTCTCTGCTGTCATGTCAGGGGTGGCGCAGAGCATCTGGCTCTCACCACTGGCTCATG 1266

QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
DB 1267 GCGTTCTTCGTGACCAAGAGTTTCAGAGCGCTTCATGGAGGTCCTCAGGCGCTTATGAGCC 1326

QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
DB 1327 TTCTGGCTTGGCTCGCTTCTGTCATCTTCACTGCTCTTTCACATTTGTCTGTCTCT 1386

QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
DB 1387 GAAACTAAAGGAAGACTCTGGAACAAATCACAGCCCATTTTGGAGGGCGCA 1437

RESULT 2
AB224792
ID AB224792 standard; cDNA; 1856 BP.
XX

1084 GCACAGCCTGTGATGCCAGCGTGGGCTGGCTGGCTGGCGGCGAGCATGTGCCTC 1143
 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuMetSerGluile 400
 1144 TTCAATCGCGGCTTTTGGCGTGGCGTGGGGGCCATCCCTCGCTCCATGTACAGATC 1203
 401 PheProLeuHisValIleGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
 1204 TTCCCTCTGCATGTCAAGGGGTGGGACAGGCATCTGGTCTTCCACCACTGCTCATG 1263
 421 AlaPheLeuValThrIleGlySerSerLeuMetGluValLeuLeuArgProTrpGlyAla 440
 1264 GCCTTTCTCGTACCAAGAGTTACAGACGCTCATGGAGTCTTTCAGGCCCTATGGAGCC 1323
 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 1324 TTCTGGCTTGGCTTTCGCTTTCGATCTTCATGTCCTTTTTCATCTTGTCTGTCTCT 1383
 461 GluThrIleGlyIleThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 1384 GAACTTAAGGAAGACTCTGGAACAAATCACAGGCCATTTTGGAGGGGCGA 1434

RESULT 3
 AAD09552
 ID AAD09552 standard; cDNA; 2080 BP.

AC AAD09552;

XX 10-SEP-2001 (first entry)

DE Human transporter and ion channel-1 (TRICH-1) cDNA.

Human; transporter and ion channel-1; TRICH-1; vaccine; cystic fibrosis;
 gene therapy; amphotropic lateral sclerosis; amnesia; muscular dystrophy;
 hypertension; angina; neurological disorder; asthma; bipolar disorder;
 dementia; depression; Alzheimer's disease; epilepsy; mood; arrhythmia;
 Pick's disease; ischaemic cerebrovascular disease; AIDS; anxiety; stroke;
 Huntington's disease; Parkinson's disease; cerebral neoplasm; allergy;
 demyelinating disease; mental disorder; schizophrenia; polymyositis;
 muscle disorder; cardiomyopathy; cataract; myocarditis; Grave's disease;
 dermatomyositis; diabetes mellitus; immunological disorder; psoriasis;
 rheumatoid arthritis; Sjogren's syndrome; systemic lupus erythematosus;
 sickle cell anaemia; Wilson's disease; infertility; Cushing's disease;
 scleroderma; pulmonary artery stenosis; neotopic; Addison's disease;
 malabsorption syndrome; hypercholesterolaemia; cancer; ss.

OS Homo sapiens.

EH Key Location/Qualifiers

FT CDS 46..1479

FT /*tag= a
 FT /product= "Human TRICH-1 protein"

XX WO200146258-A2.

XX PD 28-JUN-2001.

XX 22-DEC-2000; 2000WO-US035095.

XX 23-DEC-1999; 99US-0172000P.

XX 14-JAN-2000; 2000US-0176083P.

XX 21-JAN-2000; 2000US-0177332P.

XX 28-JAN-2000; 2000US-0178572P.

XX 02-FEB-2000; 2000US-0179758P.

XX 10-FEB-2000; 2000US-0181625P.

XX (INCY-) INCYTE GENOMICS INC.

XX Baughn MR, Burford N, Au-Young J, Lu DAM, Yang J, Reddy R;
 PI Lal P, Hillman JL, Azimzai Y, Yue H, Nguyen DB, Yao MG, Gandhi AR;
 PI Tang YT, Khan FA;

XX WPI; 2001-418042/44.

DR

P-PSDB; AA04888.

Novel human transporter and ion channel proteins useful for treating and
 preventing transport, neurological, muscle and immunological disorders.
 Claim 5; Page 140-141; 160pp; English.

The present sequence is transporter and ion channel-1 (TRICH-1) cDNA.
 TRICH is used as vaccine. TRICH is useful for treating a disease or
 condition associated with decreased expression of functional TRICH, such
 as transport disorder including amphotropic lateral sclerosis, cystic
 fibrosis, Becker's muscular dystrophy, Charcot-Marie Tooth disease,
 Duchenne muscular dystrophy, angina and hypertension, neurological
 disorders including Alzheimer's disease, amnesia, bipolar disorder,
 dementia, depression, epilepsy, ischaemic cerebrovascular disease,
 stroke, cerebral neoplasms, Pick's disease, Huntington's disease and
 Parkinson's disease, demyelinating diseases, mental disorders including
 mood, anxiety, schizophrenia and seasonal affective disorder, muscle
 disorder including cardiomyopathy, myocarditis, polymyositis,
 dermatomyositis, arrhythmias and asthma and immunological disorders
 including AIDS, adult respiratory distress syndrome (ARDS), allergies,
 anaemia, diabetes mellitus, rheumatoid arthritis, scleroderma, Sjogren's
 syndrome, systemic lupus erythematosus and other diseases including
 sickle cell anaemia, Wilson's disease, Cushing's disease, Addison's disease,
 artery stenosis, Grave's disease, hypercholesterolaemia, cancers
 glucose-galactose malabsorption syndrome, helminthic and protozoal
 psoriasis and viral, bacterial, fungal, helminthic and protozoal
 infections. TRICH DNA is useful in gene therapy and in diagnostic
 purposes

XX Sequence 2080 BP; 300 A; 688 C; 632 G; 460 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	2,14e-204	Length:	2080
Score:	2457.00	Matches:	477
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	4	Gaps:	0

US-09-886-954A-1 (1-477) x AAD09552 (1-2080)

QY	1	MetThrProGluAspProGluThrClnProLeuLeuGlyProProGlyGlySerAla	20
DB	46	ATGACGCCCGAGGACCCAGAGAAACCCAGCCGCTTCTGGGGCTCTCTGGCGGAGCGCG	105
QY	21	ProArgGlyArgArgValPheLeuAlaPheAlaAlaAlaLeuLeuGlyProLeuSerPhe	40
DB	106	CCCCCGCGCGCGGCTTCTCTCGCGCTTTCGCGCTGGCGGCCACTCAGCTTC	165
QY	41	GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAalaProPro	60
DB	166	GGCTTCGGCTCGGCTACAGCTCCCGGCCATCCCTAGCTGACGCGCGCGCGCCCCCG	225
QY	61	AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla	80
DB	226	GCCCCGCGCTGGACGACGCGCGCGCTTCTGGTTCGGGGCTGCTGCTGACCTGGGTGCC	285
QY	81	AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlySerLeu	100
DB	286	GCGCGCGGGGGAGTGTCTGGCGGCTGGGTGGACCCGCGCGCGCAAGCTGAGCTC	345
QY	101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal	120
DB	346	TTGCTGTGCTCGGCTTCGCGCTTCGCGGCTTTCGCTCATCACCGCGCGCCAGACGTG	405
QY	121	TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal	140
DB	406	TGATGCTGCTGGGGCGCGCTCTCTCACCGGCTTGGCTGGCTGTTGCTTCCCTAGTG	465
QY	141	AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys	160
DB	466	GCCCCGCTGTACATCTCCGAATCGCTACCCAGCAGTCCCGGGGTTGCTCGGCTCTCT	525

QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
DB 526 GTGAGCTAATGTCGTGCGGCATCTCTGGCTACCTGCGAGGCTGGGTGCTGGAG 585
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 586 TGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 645
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
DB 646 TTCATGCCGAGACCCGCGCTTCTGCTGACTCAGCAGAGGCCGAGAGGCCATGGCC 705
QY 221 AlaLeuArgPheLeuTrpGlySerGlnGlnGlyTrpGluAspProProIleGlyAlaGlu 240
DB 706 GCCCTGGCTTCTGTGGGCTCCGAGCAGGGCTGGGAGAGCCGCCCATCGGGCTGAG 765
QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleTyrIleGly 260
DB 766 CAGAGCTTTTACCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 825
QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
DB 826 GTCTCCCTGATGGCTTTCAGCAGCTGTGGGGTCAAGCCGTCAITTCATGAGAG 885
QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
DB 886 ACCATCTTTGAGAGGCCAAGTTCAAGGACAGCAGCTGGCTGGCTGGCTGGCTGG 945
QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuMetAspArgAlaGlyArgLeu 320
DB 946 ATCCAGGTGCTGTTCAGCTGTGGGGCTCTCATCATGACAGCAGGCGGAGGCTG 1005
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
DB 1006 CTCTGGCTTTGTCAGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGTGGTGT 1065
QY 341 LysLeuThrGlnGlyProGlyAenSerSerHisValAlaIleSerAlaProValSer 360
DB 1066 AAGCTGACCCAGGCTGGCTGGCAACTCTGCGACGTGGCCATCTCGGGCTGTCT 1125
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
DB 1126 GCACAGCTGTGATGCCAGCGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1185
QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProIleProIleMetSerGluIle 400
DB 1186 TTCATGCCGCTTTGCGTGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGGGCTGG 1245
QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
DB 1246 TTCCTCTGTCATGTCAAGGGCTGGGCGACAGGCATCTGGCTCTCAACCACTGGCT 1305
QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
DB 1306 GCCTTTCTGTCACCAAGAGTTTCAGCAGCTCATGGAGTCTCATGGGCTCATGGAG 1365
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
DB 1366 TTCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1425
QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
DB 1426 GAAACTAAGGAAGAACTCTGGACAACTACAGCCCATTTTGGAGGGCGA 1476

RESULT 4

ID ABZ24793
XX standard; cDNA; 1873 BP.
AC ABZ24793;
XX
DT 07-APR-2003 (first entry)
XX

DE Human solute carrier type 2A nucleic acid 7657680.
XX Human; solute carrier type 2A; SLC2A; glucose transporter; p53; cancer;
KW cytostatic; gene therapy; gene; ss.
XX Homo sapiens.
XX Key Location/Qualifiers
FH 4..1437
FT /*tag= a
TX /product= "Human SCL2A"
XX WO200298467-A1.
XX 12-DEC-2002.
XX 03-JUN-2002; 2002WO-US017419.
XX 05-JUN-2001; 2001US-0296076P.
PR 10-OCT-2001; 2001US-0328605P.
PR 15-FEB-2002; 2002US-0357253P.
XX (EXEL-) EXELIXIS INC.
XX Friedman L, Florman GD, Belvin M, Francis-Lang H, Li D, Funke RP;
PI WPI: 2003-201283/19.
XX P-PSDB; ABP58364.
XX Identifying candidate p53 pathway modulating agent as therapeutic target
PT for disorders of defective p53 function e.g. cancer, by assaying purified
PT solute carrier family 2 (SLC2A) polypeptide or nucleic acid with a test
PT agent.
XX Disclosure; Page 43-44; 58pp; English.
XX The present sequence is that of human solute carrier type 2A (SLC2A)
CC nucleic acids 7657680. Genetic screens were designed to identify
CC modifiers of the p53 pathway in Drosophila in which p53 was
CC overexpressed. Human orthologues (polynucleotides and polypeptides) of
CC one such modifier were then identified, including nucleic acid 7657680.
CC SLC2As are glucose transporter proteins with sugar transporter domains.
CC SLC2A nucleic acids and polypeptides are attractive drug targets for the
CC treatment of pathologies associated with a defective p53 signalling
CC pathway, such as cancer. The invention provides in vitro and in vivo
CC methods of assessing SLC2A function. Modulation of an SLC2A or its
CC binding partners is useful for understanding the association of the p53
CC pathway and its members in normal and disease conditions and for
CC developing diagnostic and therapeutic modalities of p53-related
CC pathologies. SLC2A-modulating agents that act by inhibiting or enhancing
CC SLC2A expression, directly or indirectly, e.g. by affecting an SLC2A
CC function such as transport or binding activity, can be identified using
CC methods provided. Modulators include small molecules, nucleic acids,
CC antibodies, antisense oligonucleotides and phosphothioate morpholino
CC oligomers (claimed)
XX Sequence 1873 BP; 279 A; 621 C; 573 G; 400 T; 0 U; 0 Other;
SQ Alignment Scores:
Pred. No.: 1.14e-203 Length: 1873
Score: 2448.00 Matches: 475
Percent Similarity: 99.79% Conservative: 1
Best Local Similarity: 99.58% Mismatches: 1
Query Match: 99.63% Indels: 0
DB: 8 Gaps: 0
US-09-886-954A-1 (1-477) x ABZ24793 (1-1873)
QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
DB 4 ATGACGCCCGAGACCCAGAGAAACCCAGCCGCTTCTGGGGCTCTCTGGGCGAGCGG 63
QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40

Db	64	CCCCGGCGCGCGTCTTCTGCGCGCTTGGCGCGCTGCGCGCGCTCAGCTTC	123	Db	1144	TTTCATCGCGCGCTTTGCGGTGGGGCCCATCCCTCGGTCTCATGTGAGATC	1203
Qy	41	GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro	60	Qy	401	PheProLeuHisVallysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet	420
Db	124	GCGTTTGGCGTGGTACAGTCCCGCGCATCCCTAGCCTGCGCGCGCGCGCGCGCG	183	Db	1204	TTCCCTCTGATGTCATGAAGGGCGTGGCGAGCAGCATCTGCGTCTCACCACCTGGCTCATG	1263
Qy	61	AlaProArgLeuAspAlaAlaSerTrpPheGlyAlaValValTrpLeuGlyAla	80	Qy	421	AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla	440
Db	184	GCCCGCGCGTGGACGCGCGCGCTTCTGTTTGGGGTGTGTCGACCTGGGTGCC	243	Db	1264	GCTTTCTTCTGTGACCAAGAGATTACAGCCTCTGAGAGTCTCAGGCGCTTATGGAGCC	1323
Qy	81	AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu	100	Qy	441	PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro	460
Db	244	GCGCGCGCGGAGTGTGGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	303	Db	1324	TTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT	1383
Qy	101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal	120	Qy	461	GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg	477
Db	304	TGCTGTGCTGCGTGGCTTGTGGCGGCTTGGCGTCAACCGCGCGCGCGCGCGCG	363	Db	1384	GAAATTAAGGAAGAACTCTGGAAACAATCACAGCCCATTTTGGGGGCGA	1434
Qy	121	TrpMetLeuLeuGlyGlyArgLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal	140	RESULT 5			
Db	364	TGGATGCTGCTGGGGGCGCGCTCTCTACCGGCTGGCTGGCTGGCTGGCTGGCTGG	423	ID	AAF55865	standard; cDNA; 2217 BP.	
Qy	141	AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys	160	XX	AAF55865;		
Db	424	GCCCCGCTTACATCTCCGAAATCGCTTACCCAGCAGTCCGGGGTGTGCTGGCTTCTGT	483	XX	17-APR-2001	(first entry)	
Qy	161	ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu	180	DE	Human	GLUTX1 coding sequence.	
Db	484	GTCAGCTAATGTCTGTGGCATCTCTCTGGCTTACCTGGCTTACCTGGAGGCTGGTGGAG	543	XX	Human;	GLUTX; gene therapy; vaccine; hexose transport modulator;	
Qy	181	TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys	200	KW	hexose transport disorder; ischaemia; diabetes; hypoglycaemia; ss;		
Db	544	TGGCGCTTGGCTGTGCTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	603	XX	hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.		
Qy	201	PheMetProGlnThrProArgPheLeuThrGlnHisArgArgGlnGlnAlaMetAla	220	OS	Homo sapiens.		
Db	604	TTATGCCCCGAGACCCCGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	663	XX	WO200104145-A2.		
Qy	221	AlaLeuArgPheLeuTrpGlySerGlnGlnGlyTrpGluAspProIleGlyAlaGlu	240	XX	18-JAN-2001.		
Db	664	GCCCTGCGTCTCTGTGGGCTCCGAGAGGGTGGGAAGACCCCGCTGCGGGGTGAG	723	XX	14-JUL-2000; 2000WO-IB001042.		
Qy	241	GlnSerPheHisLeuAlaLeuArgGlnProGlyIleTyrLysProPheIleLeGly	260	XX	14-JUL-1999; 99US-0143907P.		
Db	724	CAGAGCTTTTACCTGGCGCTGCTGCGGACGCCCGCATCTACAGCCCTTTCATCGGC	783	PR	27-AUG-1999; 99US-0151140P.		
Qy	261	ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu	280	PR	23-FEB-2000; 2000US-0184285P.		
Db	784	GTCTCCCTGATGGCTTCCAGCAGCTGTCGGGGTCAACCGCTCATGTTCATGACAG	843	PR	13-JUL-2000; 2000US-00616132.		
Qy	281	ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal	300	XX	(UYLA-) UNIV LAUSANNE.		
Db	844	ACCATCTTTGAAGAGGCAAGTTCAAGGACAGCAGCCTGGCCTGGGTCGTCGGGTGC	903	XX	Thorens B, Ibberson M, Uldry M;		
Qy	301	IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu	320	XX	WPI; 2001-112615/12.		
Db	904	ATCCAGGTGCTTTTACAGCTGTGGCGCTCTCATCATGTGACAGAGCAGGGCGAGCTG	963	DR	P-PSDB; AAB66932.		
Qy	321	LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe	340	XX	Nucleic acids encoding GLUTX glucose transporter proteins, useful in the		
Db	964	CTCCTGGTGTGTGAGGTGTGTGATGTGTTCAGCAGCAGTGTGCTTTCGGCGCTTCTTC	1023	PT	prevention, diagnosis and treatment of hexose transport disorders, e.g.		
Qy	341	LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer	360	PT	ischemia and diabetes.		
Db	1024	AAGCTGACCCAGGGTGGCGCTTGGCAACTCTCTGCGACGTGGCCATCTCGGGCGCTCTCT	1083	XX	Claim 3; Page 70-71; 124pp; English.		
Qy	361	AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu	380	PS	The present invention relates to GLUTX proteins (AAF55865-AAF55871 and		
Db	1084	GCACAGCTGTGTATGCCAGCGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTG	1143	CC	AAB66932-AAB66941). The GLUTX proteins are related to the facilitative		
Qy	381	PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuMetSerGluLe	400	CC	glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport		
				CC	function. The GLUTX proteins may be used in the diagnosis, prevention and		
				CC	treatment of hexose transport disorders such as ischaemia, diabetes,		
				CC	hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a		
				CC	neurodegenerative disease. The present sequence is the coding sequence		
				CC	for human GLUTX1		
				XX	Sequence 2217 BP; 320 A; 750 C; 672 G; 475 T; 0 U; 0 Other;		
				SQ			
				Alignment Scores:			
				Pred. No.:	1.42e-203	Length:	2217
				Score:	2448.00	Matches:	475
				Percent Similarity:	99.79%	Conservative:	1

Best Local Similarity:	99.58%	Mismatches:	1
Query Match:	99.63%	Indels:	0
DB:	4	Gaps:	0
US-09-886-954A-1 (1-477) x AAF55865 (1-2217)			
QY	1	MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla	20
DB	348	ATGACGCCCGGAGACCCAGAGGAAACCCAGCCGCTTCTGGGGCTCTCTGGCGGCGCGG	407
QY	21	ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe	40
DB	408	CCCCGGCGCGCGCGCTTCTTCGCGCGCTTCTGGCGCTTCTGGCGCGCGCTTC	467
QY	41	GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro	60
DB	468	GGCTTCGGCTCGGCTACAGTCCCGCGCATCCCTAGCTGACGGCGCGCGCGCGCG	527
QY	61	AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla	80
DB	528	GDCCCGCGCTGACGACGCGCGCTCTCTGGTTCGGGCTGCTGACCCCTGGGTGCC	587
QY	81	AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLeuSerLeu	100
DB	588	GGCGCGGGGGAGTGTCTGGCGGCTGGCTGGTGGACCGCGCGCGCGCAAGCTGAGCCTC	647
QY	101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal	120
DB	648	TTGCTGTCTCCGTGCGCTTCTGGTGGCGGCTTTCCTGCTCATCCCGCGCGCGCGAGCGTG	707
QY	121	TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal	140
DB	708	TGGATGCTGTGGGGGGCGCTCTCACCGGCTGGCGCTGGCGGTGTGCTCCCTAGTG	767
QY	141	AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys	160
DB	768	GCCCGGCTTACATCTCCGAATCGCTACCCAGCAGTCCGGGGGTGTCTCGGCTCTGT	827
QY	161	ValGlnLeuMetValValValGlyLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu	180
DB	828	GTCACGTAATGTCTGTGGCATCTCTCGGCTTACCTGGCAGCTGGGTGGGTGGAG	887
QY	181	TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys	200
DB	888	TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	947
QY	201	PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla	220
DB	948	TTTCATGCCGAGACCCCGGCTTCTCTGCTGACTCAGCAGCGCGCGCGCATGGCC	1007
QY	221	AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheGlyAlaGlu	240
DB	1008	GCCTCGGGTCTGTGGGGCTCCGAGAGGGCTGGAGAGCCCGCCATCGGGGTGAG	1067
QY	241	GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIlePhePheLeuGly	260
DB	1068	CAGAGCTTTCACCTGGCCCTGCTGGCGCGCGCGCATCTACAAGCCCTTCATCGGC	1127
QY	261	ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu	280
DB	1128	GTCTCCCTGATGGCTTCCAGCAGCTGTCTGGGGGTCAACCGCGCTCATGTTCTATG	1187
QY	281	ThrIlePheGluGluAlaPheLeuAspSerSerLeuAlaSerValValValGlyVal	300
DB	1188	ACCATCTTTGAAGAGCGCAAGTTCAAGCAGCAGCGCTGGCTGGGTGGGTGGTGC	1247
QY	301	IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLeu	320
DB	1248	ATCCAGGTGCTGTTCACAGCTGTGGCGCTCTCATATGACAGCAGCGCGCGAGCTG	1307
QY	321	LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe	340
DB	1308	CTCCTGGTCTTGTGAGGTGGTGTGATGGTGTTCAGCAGCAGTGCCTTCGCGCGCTACTTC	1367

QY	341	LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer	360
DB	1368	AACTGACCCAGGGTGGCTTGGCAACTCTCTGCACGTGGCCATCTCGGGCGCTGTCTCT	1427
QY	361	AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu	380
DB	1428	GCACAGCTGTGATGCCAGCGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG	1487
QY	381	PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuMetSerGluIle	400
DB	1488	TTTCATCGCGCGCTTTCGGGTGGGCTGGGCGCCCATCCCTGGCTCTCATGTGAGATC	1547
QY	401	PheProLeuHisValLysGlyValAlaThrGlyVileCysValLeuThrAsnTrpLeuMet	420
DB	1548	TTCCCTCTGCATGTCAAGGGCTGGGCGACAGCATCTGGCTCTCACCAACTGGCTCATG	1607
QY	421	AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla	440
DB	1608	GCCTTTCTGTCACCAAGAGTTTCAGCAGCTCATGGAGTCTTCAGGCGCTATGGAGCC	1667
QY	441	PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro	460
DB	1668	TTCTGGCTTGGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTCGCTTTC	1727
QY	461	GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg	477
DB	1728	GAAATTAAGGAAGAGTCTGGAACAATCACAGCCATTTTGAGGGCGGA	1778
RESULT 6			
ID	AAF55866		
XX	AAF55866 standard; cDNA; 2087 BP.		
AC	AAF55866;		
XX			
DT	17-APR-2001 (first entry)		
XX			
DE	Rat GLUTX1 coding sequence.		
XX			
KW	Rat; GLUTX; gene therapy; vaccine; hexose transport modulator;		
KW	hexose transport disorder; ischaemia; diabetes; hyperglycaemia; ss;		
KW	hypoglycaemia; glucose metabolism disorder; neurodegenerative disease.		
XX			
OS	Rattus sp.		
XX			
PN	WO200104145-A2.		
XX			
PD	18-JAN-2001.		
XX			
PF	14-JUL-2000; 2000WO-IB001042.		
XX			
PR	14-JUL-1999; 99US-0143907P.		
PR	27-AUG-1999; 99US-0151140P.		
PR	23-FEB-2000; 2000US-0184285P.		
PR	13-JUL-2000; 2000US-00616132.		
XX			
PA	(UYLA-) UNIV LAUSANNE.		
XX			
PI	Thorens B, Ibberson M, Uldry M;		
XX			
XX	WPI; 2001-112615/12.		
DR	P-PSDB; AAB66933.		
XX			
PT	Nucleic acids encoding GLUTX glucose transporter proteins, useful in the		
PT	prevention, diagnosis and treatment of hexose transport disorders, e.g.		
PT	ischaemia and diabetes.		
XX			
PS	Claim 3; Page 71-73; 124pp; English.		
XX			
CC	The present invention relates to GLUTX proteins (AAF55865-AAF55871 and		
CC	AAB6932-AAB6941). The GLUTX proteins are related to the facultative		
CC	glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport		
CC	function. The GLUTX proteins may be used in the diagnosis, prevention and		

DR WPI; Z001-112015/
DR P-PSDB; AAB66934.

4

XX Nucleic acids encoding GLUTX glucose transporter proteins, useful in the
 PT prevention, diagnosis and treatment of hexose transport disorders, e.g.
 PT ischemia and diabetes.

XX Claim 3; Page 73-74; 124pp; English.

XX The present invention relates to GLUTX proteins (AAF55865-AAF55871 and
 CC AAB66932-AAB66941). The GLUTX proteins are related to the facultative
 CC glucose carriers GLUT1-GLUT5 and have hexose binding and/or transport
 CC function. The GLUTX proteins may be used in the diagnosis, prevention and
 CC treatment of hexose transport disorders such as ischaemia, diabetes,
 CC hyperglycaemia, hypoglycaemia, a glucose metabolism disorder and/or a
 CC neurodegenerative disease. The present sequence is the coding sequence
 CC for murine GLUTX1

XX SQ Sequence 2072 BP; 351 A; 673 C; 584 G; 464 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.33e-177 Length: 2072
 Score: 2150.00 Matches: 409
 Percent Similarity: 92.47% Conservative: 33
 Best Local Similarity: 85.56% Mismatches: 34
 Query Match: 87.51% Indels: 2
 DB: 4 Gaps: 2

US-09-886-954A-1 (1-477) x AAF55867 (1-2072)

QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 DB 21 ARGTCCTCCAGAGACCCAGGAGAGCGCCGCTATTGCGGCACCCGGAAGCCAGGACT 80
 QY 21 ProArgGlyArgArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 DB 81 CQCGCGCGCGCGGCTCTCTCGCTTCCTTCGCGCGCGCTGCGGACCCCTCAACTTC 140
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaProPro 60
 DB 141 GGCCTTCGCGCTCGCTGCTACAGCTCCCGCGCATCCCGAGCTGCGGCGCACCGCCCG 200
 QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 DB 201 GGCCTCGGCTCGAGACAAATGCGGCTCTGCTTCGCGGCGCTGCTGACCTTGGGCGCT 260
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
 DB 261 GCTGAGGGGGGATACCTGCGGCGCTGCTCTCGGACCGTTCAGGGGCGAAGCTGAGCTTC 320
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 DB 321 TTGCTCTGCACCGTGGCTTCGCTGACTGGCTTTGCTGTCTCATCACCGCGCGCGGATGTG 380
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 DB 381 TGGATGCTGCTCGAGCGCGCTCTCTCACCGGCTTGGCTGCGGAGTCCGCTCCTACCTAGT 440
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 DB 441 GCACCGGTTTACATCTCGGAATCGCTACCCAGCTGTCCGAGGACTGCTCGGCTCTCTGT 500
 QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 DB 501 GTGAGCTGATGTTGTCTCATCGCATCTCTGCGCTATGTGCGAGGCTGGGTCTCTAGAG 560
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 DB 561 TGGCGCTGCTGCGCGCTGTGCGGCTGTGCGGCTGCTGCGGCTGCTGCTGCTGCTGCTGCT 620
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnAlaMetAla 220
 DB 621 TACATGCGCGAGACCCAGTTTCTCTCATCTCAACACAGTACACAGGAGGCGGCTGGCT 680
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheGlyAlaGlu 240

DB 681 GCCTTCGCGCTCTCTGCGGCTCTGAGGAGGCTGGGAAGAGAGCCCTCTGTTGGGCTGAG 740
 QY 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleTyrIlePheIle 259
 DB 741 CACCAAGGGCTTCAGCTGGCCCTGCTGAGGCGCCCTGGGCATCTACAAAGCCCTCATCATC 800
 QY 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAla 279
 DB 801 GGCATTTCCTCATGCTCTTCAGCAGCTGTTCAGGCGTCAATGCTATCATGTTCTATGCC 860
 QY 280 GluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGly 299
 DB 861 AACAGCATCTTCAGGAGGCGCAAGTTCAAGGACAGCAGCTGCGCTCGGTCACTGTGGGC 920
 QY 300 ValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArg 319
 DB 921 ATAATCCAGGCTCTGTTCACTGCTGCGGCGCCCTCATCATGACAGACAGAGCGCGAAGG 980
 QY 320 LeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
 DB 981 CTGCTCTCGGCTTGTGCGGCTGATCATGGTGTTCAGTATGAGTGCCTTGGTACTTAC 1040
 QY 340 PheLysLeuThrGlnGlyGlyProGlyAenSerSerHisValAlaIleSerAlaProVal 359
 DB 1041 TTCAAACTGACCCAGAGCCTCCCGACCACTCTCTCCACGTAGGCGCTG---GTGCCCATC 1097
 QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCys 379
 DB 1098 GCGGCGGAGCCCTGTGGATGTCCAAGTGGGACTGGGCTGGCTGGCTGGCTGGCTGGCTGG 1157
 QY 380 LeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGlu 399
 DB 1158 CTCTTCATGCTGGCTTTGCGGTGGCTGGGAGCCCATCCCTGGCTCTCATGTGACAG 1217
 QY 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAenTrpLeu 419
 DB 1218 ATCTTCTCTCTGATGTCAAGGGTGTGGCTACCGCATCTGTGTCTCTCAACCACTGGTTC 1277
 QY 420 MetAlaPheLeuValThrLysGlyPheSerSerLeuMetGluValLeuArgProTyrGly 439
 DB 1278 ATGGGCTTCTAGTACCAAGAGTTCAACAGGTCATCGAGATGCTCAGACCTACCGT 1337
 QY 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
 DB 1338 GCCTTCGCTCACCGCTGCTCTGCGCTCTCAGTGTCTTATTCACACTGACCGTGTTC 1397
 QY 460 ProGluThrLysGlyLysThrLeuGluIleThrAlaHisPheGluGlyArg 477
 DB 1398 CCTGAGACTAAAGGAGGAGGACTCTGGAACAAGTCAAGGAGGAGGAGGAGGAGGAGG 1451

RESULT 8

ABX34029
 ID ABX34029 standard; cDNA; 1954 BP.

AC ABX34029;

XX 10-FEB-2003 (first entry)

DE Human cancer suppressing protein PF7425.

XX Human; gene; ss; cancer suppressing protein; cancer.

XX Homo sapiens.

XX CN1351081-A.

XX 29-MAY-2002.

XX 31-OCT-2000; 2000CN-00127102.

XX 31-OCT-2000; 2000CN-00127102.

XX XX

200 sPheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAl 22

XX
CE-MAR-2004

XX XX

XX 12-SEP-2003; 2003WO-US028227.
XX 12-SEP-2002; 2002US-0410259P.
XX 12-SEP-2002; 2002US-0410260P.
XX (INCY-) INCYTE CORP.
XX Schmidt JP, Wright RJ, Bruns CM, Marjanovic MM, Shen F, Harthehorn TA, Suchorolski MT, Altus CM, Pitts SJ, Elder LV, Mooney EM, Delegeane AM, Panesar IS, Banville SC, Reddy TP, Stevens KA, Blanchard JL, Panzer SR, Wang X, Au AP, Gerstin EH, Peralta CH, Anderson SB, Rioux P, Shen EJ, Wu MC, Stuve LL, Lagace RE, Spiro PA, Stewart EA, Wingrove J, Vitt UA, Kirton BS, Xu Y, Kwong M, Policky JL, Hurwitz BL, Ma Y, Jackson JL, Gietzen D, Patury S, Shi X, Suarez CJ;
XX WPI: 2004-329368/30.
XX P-PSDB; ABW4862.
XX New diagnostic and therapeutic polynucleotides and polypeptides, useful in diagnosing a condition, disease or disorder associated with human molecules, e.g. autoimmune or inflammatory disorders, in gene therapy or in gene mapping.
XX Claim 1; Page; 190pp; English.
XX The invention relates to novel diagnostic and therapeutic polynucleotides selected from one of the 2722 sequences defined in the specification. A polynucleotide of the invention may have a use in gene therapy. The human diagnostic and therapeutic polynucleotides (dithp) or polypeptides may be used to diagnose a particular condition, disease or disorder associated with human molecules, e.g. cell proliferative disorders, autoimmune/inflammatory disorder, developmental disorder, endocrine disorder, neurological disorders, gastrointestinal disorders, or infections caused by virus, bacteria, fungi or parasite. The dithp molecules may also be used in genetic mapping, in identifying individuals from minute biological samples, in detecting single nucleotide polymorphisms, as molecular weight markers, and for somatic or germline gene therapy. The present sequence represents a dithp polynucleotide of the invention. Note: The sequence data for this patent is not represented in the printed specification, but was obtained in electronic format directly from WIPO at www.wipo.int/pct/en/sequences/listing.htm
XX Sequence 1923 BP; 286 A; 651 C; 575 G; 411 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 1,91e-170 Length: 1923
Score: 2067.50 Matches: 408
Percent Similarity: 85.53% Conservative: 0
Best Local Similarity: 85.53% Mismatches: 0
Query Match: 84.15% Indels: 69
DB: 13 Gaps: 1
US-09-886-954A-1 (1-477) x ACN43514 (1-1923)
QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
DB 268 ATGACGCCCGAGGACCCAGAGGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGACGCGG 327
QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
DB 328 CCCCAGGCGCGCGCGCTCTCTCGCGCGCTTCTCGCGCTTCTCGCGCTTCTCGCGCTTCT 387
QY 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
DB 388 GGCCTTCGCGCTCGGCTACAGCTCCCGGCGCATCCCTAGGCTGCGAGCGCGCGCGCGCG 447
QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
DB 448 GCCCGCGCGCTGGAGCAGCGCGCGCGCTCTCTGGTTTCGGG----- 486
QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100

DB 486 ----- 486
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
DB 486 ----- 486
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
DB 486 ----- 486
QY 141 AlaProValTyrIleSerGlnIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
DB 487 -----GTCTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGTGTCTCGGCTCTCTGT 540
QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
DB 541 GTGCAGCTAATGCTGCTCGCATCTCTCGGCTTCTGCGAGGCTTGGGCTGGGCTGGAG 600
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 601 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 660
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
DB 661 TTCTGCGCGAGACCCCGCGCTTCTGCTGACTCAGCAGCAGCGCGCGGCGCATGGCC 720
QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
DB 721 GCCCTGCGGTTCTGTGGGGCTCCGAGCAGGGCTGGGAGACCCCCCATCGGGGCTGAG 780
QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleLeuGly 260
DB 781 CAGAGCTTTCACCTGGCCCTGTCTGGCGAGCCCGCATCTACAGCCCTTTCATCATCGGC 840
QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
DB 841 GTCTCCCTCATGCGCTTCCAGCAGCTGTCGGGGGTCAAGCCGCTCATGTTCATGACAG 900
QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyVal 300
DB 901 ACCATCTTTGAAGAGCCCAAGTTCAAGACAGAGCGCTGGCTCGGTCGTGGGGGTGC 960
QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
DB 961 ATCCAGGTGCTGTTACAGCTGTGGCGCTCTCATCATGACAGCAGCGCGGCGAGGCTG 1020
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
DB 1021 CTCTCTGGTCTTCTCAGGTGTGCTCATGTGTTCAGCAGCAGTGCCTTCGCGCGCTACTTC 1080
QY 341 LysLeuThrGlnGlyProGlyLysSerSerHisValAlaIleSerAlaProValSer 360
DB 1081 AAGCTGACCCAGGGTGGCCCTGGCAACTCTCTCGAGTGGCCATCTCTGGCGCGCTGTCTCT 1140
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
DB 1141 GCACAGCTGTGATGCCAGCGTGGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 1200
QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
DB 1201 TTCTATCGCGCGCTTTCGCGTGGGCGCCATCCCTGGCTCTCATGTGACAGATC 1260
QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrLeuTrpLeuMet 420
DB 1261 TTCCCTCTGCTGTCAGGTGAGGGCGTGGCGACAGGCATCTGCGCTCTCAACCTGGCTCATG 1320
QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
DB 1321 GCCTTCTGTCGACCAAGAGTTTCAGAGCTTCATGGAGTCTCTAGGGCCCTATGGAGCC 1380
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460

Percent Similarity:	80.08%	Conservative:	0
Best Local Similarity:	80.08%	Mismatches:	0
Query Match:	77.84%	Indels:	95
DB:	13	Gaps:	2

QY	1	MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla	20
DB	268	ATGACGCCCGAGGACCCAGAGAAACCCAGCCGCTCTGGGGGCTCTCTGGCGGACGCG	327
QY	21	ProArgGlyArgArgValpheLeuAlaAphaAlaAalaLeuGlyProLeuSerPhe	40
DB	328	CCCCCGGGCGCGCGTCTTCTGCGCGCTTCGCCCTGCCCTGGGCCACTCAGCTTC	387
QY	41	GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnhrGAlaAProPro	60
DB	388	GGCTTCGCGCTCGGTACAGCTCCCGGCCATCCCTAGCCTGAGCGCGCGCGCCCCG	447

[illegible]

Db	486	-----	486
Qy	141	AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys	160
Db	487	-----GTCACATCTCCGAATACGCCTACCCAGCAGTCCGGGGTGTCTCGGCTCCTGT	540
Qy	161	ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu	180
Db	541	GTGCAGCTAATGGTCGTCGTGGGCATCTCTCTGGCCTACCTGGCAGGCTGGTGCTCGAG	600
Qy	181	TyrArgTyrLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys	200
Db	601	TGGCGCTGGCTGGCTGTGTGGCTGCGTGGCCCCCTCTCCTCATGCTGCTTCTCATGTGC	660
Qy	201	PheMetProGlnThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAla	220
Qy	221	TATATTCACCAATGACCGCGCGCTTCTGTGTAATTCAGCACAGCGCCACAGAGCCATGATGCC	240

[illegible]

Qy 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 34

```
Db 943 CTTCTGCTCTTGTGTCAGGTGGTTCATGGTGTTCAGCAGAGTGCCTTCGGCGCTACTTC 1002
QY LysLeuThrGlnGlyProGlyAenSerHisValAlaIleSerAlaProValSer 360
Db 1003 AAGCTGACCCAGGGTGGCCCTGGCACTCTTCGACGTGGCCATCTCGGCGCTGTCTCT 1062
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
Db 1063 GCACAGCCTCTTGATGCCAGGTGGGCTGGCCCTGGCTGGCGGTGGCGACATGTGCCTC 1122
QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
Db 1123 TTCATGCCCGCTTTCGGTGGGCTGGGGCCCATCCCTGGCTCTCATGTTCAGAGATC 1182
QY 401 PheProLeuHisValIysGlyValAlaThrGlyIleCysValLeuThrAenTrpLeuMet 420
Db 1183 TTCCTCTGATGTCAAGGGCTGGCGACAGGATCTGCGTCTCCAACTGGCTCATG 1242
QY 421 AlaPheLeuValThrIysGluPheSerSerLeuMetGluValLeuArgProTrpGlyAla 440
Db 1243 GCCTTTCTCGTACCAAGAGTTTCAGCAGCCTCATGGAGGTCTTCAGGCCCTATGGAGCC 1302
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
Db 1303 TTCGGCTTGCCTCCGCTTTCTGCATCTTCAGTGTCTTTTCACCTTTGTCTGTCCCT 1362
QY 461 GluThrIysGlyIysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1363 GAACTAAGGAAGACTCTGGAAACAATCACAGCCCATTTTGAAGGGCGA 1413
RESULT 11
ADL33401
ID ADL33401 standard; DNA; 1790 BP.
XX
AC ADL33401;
XX
DT 20-MAY-2004 (first entry)
XX
DE Human transporter and ion channel (TRICH) gene #46.
XX
KW anti-HIV; anti-allergic; anti-inflammatory; antianemic; antiparkinsonian;
KW neotropic; anticonvulsant; antiarteriosclerotic; antiasthmatic;
KW immunosuppressive; antihypertoid; cytosstatic; hepatotropic; dermatological;
KW antidiabetic; aephrotropic; antigout; thyromimetic; neuroprotective;
KW osteopathic; antiarthritic; antiparasitic; antihelminthic; antipsoriatic;
KW uropathic; ophthalmological; antirheumatic; hemostatic; antibacterial;
KW virucide; protozoacide; fungicide; gene therapy; gene; ds.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT allele replace(826,T)
FT /tag= a
FT allele replace(827,T)
FT /tag= b
FT allele replace(870,G)
FT /tag= c
FT allele replace(871,G)
FT /tag= d
FT allele replace(1274,A)
FT /tag= e
FT allele replace(1506,G)
FT /tag= f
FT allele replace(1507,G)
FT /tag= g
FT allele replace(1541,T)
FT /tag= h
FT allele replace(1582,G)
FT /tag= i
FT allele replace(1617,T)
FT /tag= j
XX
PN W02003083085-A2.
```

```
XX 09-OCT-2003.
XX 27-MAR-2003; 2003WO-US0093797.
XX 28-MAR-2002; 2002US-0368840P.
XX 26-APR-2002; 2002US-0375637P.
XX (INCY-) INCYTE CORP.
XX Marquis JP, Lee SY, Emerling BM, Hafalia AJA, Khare R, Kable AE;
XX Richardson TW, Swarnakar A, Chawla NK, Becha SD, Mason PM;
XX Elliott VS, Ramkumar J, Griffin JA, Tran UK, Ison CH, Lindquist EA;
XX Jiang X, Jackson AA, Wilson AD, Jin P, Chang H;
XX WPI; 2003-833535/77.
XX P-PSDB; ADL33342.
XX New human transporters and ion channels (TRICH) and polynucleotides,
XX useful for diagnosing, treating or preventing autoimmune or inflammatory
XX disorders (e.g. AIDS, allergy or anemia), multiple sclerosis, cancer or
XX hepatitis.
XX Claim 12; SEQ ID NO 105; 405pp; English.
XX The invention relates to an isolated polypeptide (I), which is a human
XX intracellular signaling molecule, which is a human intracellular
XX signaling molecule, a naturally occurring amino acid sequence at least 90%
XX -96% identical to it or a biologically active fragment or an immunogenic
XX fragment of the polypeptide. The human TRICH, polynucleotides, agonists
XX and antagonists are useful for diagnosing, treating or preventing
XX disorders associated with aberrant expression of TRICH, particularly cell
XX proliferative disorders (e.g. arteriosclerosis, atherosclerosis,
XX cirrhosis, hepatitis, paroxysmal nocturnal hemoglobinuria, polycythemia
XX vera, psoriasis, primary thrombocytopenia or cancer), developmental
XX disorders (e.g. renal tubular acidosis, anemia or mental retardation),
XX neurological disorders (e.g. Alzheimers disease, Parkinson's disease or
XX epilepsy), autoimmune/inflammatory disorders (e.g. AIDS, allergies,
XX asthma, autoimmune thyroiditis, contact dermatitis, Crohn's disease,
XX diabetes mellitus, glomerulonephritis, Goodpasture's syndrome, gout,
XX Graves' disease, Hashimoto's thyroiditis, irritable bowel syndrome,
XX multiple sclerosis, osteoarthritis, osteoporosis, pancreatitis, Reiter's
XX syndrome, rheumatoid arthritis, Sjogren's syndrome, uveitis), or viral,
XX bacterial, fungal, parasitic, protozoan or helminthic infections. The
XX polynucleotides encoding TRICH are useful for creating transgenic animals
XX to model human disease. This sequence corresponds to the gene encoding
XX one of the proteins of the invention.
XX SQ Sequence 1790 BP; 279 A; 587 C; 535 G; 389 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3 42e-149 Length: 1790
Score: 1823.50 Matches: 370
Percent Similarity: 77.41% Conservative: 0
Best Local Similarity: 77.41% Mismatches: 0
Query Match: 74.22% Indels: 108
DB: 11 Gaps: 1
US-09-886-954A-1 (1-477) x ADL33401 (1-1790)
QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
Db 47 ATGACGCCCGGAGGACCCAGAGAAACCCAGCCGCTTCGGGGCTCTCTGGCGCAGCGCG 106
QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db 107 CCGCGCGCGCGCGCTCTTCCTCGCGCTTCGGCGCTTCGGCGCTTCAGCTTC 166
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaProPro 60
Db 167 GGTTCGGCTCGGCTACAGTCCCGGCCCATCCCTAGCTTCGAGCGCGCGCGCCCCG 226
QY 61 AlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValThrLeuGlyAla 80
```


Db	227	GCCTCGCGCTGGACGCGCGCGCTCTCGGGTTCGTCGACCTCGGTGCC	286	Db	1029	-----CTCAGGCCCTATGGAGC	1045
Qy	81	AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu	100	Qy	440	aPheTrpLeuAlaSerAlaPheCysAlaPheSerValLeuPheThrLeuPheCysValPr	460
Db	287	CGCGCGGGGGAGTCTGGGGGCTGGCTGGTGACCGCGCGGGCGGCAAGCTGAGGCTC	346	Db	1046	CTTCTGGCTTGGCTCGCTTCTGCATCTTCAGTGTCCCTTTTCACCTTTTCTGTGTGCC	1105
Qy	101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal	120	Qy	460	oGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg	477
Db	347	TTCCTGTGCTCGTCCCTTCTGTGGCGGCTTTCGCTCATCACCGCGGCCACGAGCTG	406	Db	1106	TGAACCTAAGGAAGACCTCTGGACCAATCACAGCCCAATTTTGAGGGGCGA	1157
Qy	121	TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal	140	RESULT 12			
Db	407	TGATGCTCTGGGGGCGCGCTCTCCACCGGCTGGCTGGCTGGCTGGCTGGCTGGCT	466	ID	ADH45466	standard; DNA; 1763 BP.	
Qy	141	AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuGlySerCys	160	XX	ADH45466;		
Db	467	GCCCCGGTCTATCTCCGAAATCGCTACCCAGCAGTCCGGGGTTCCTCGGCTCTGT	526	AC	25-MAR-2004	(first entry)	
Qy	161	ValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu	180	XX	Human molecule useful for disease detection + treatment DNA, SEQ ID 60.		
Db	527	GTGACGCTAATGTGCTGCTGGCATCTCTCTGGCTACCTGGCAGGCTGGTGTGGAG	586	XX	human; molecule; disease detection; treatment; MDDT; cytostatic;		
Qy	181	TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys	200	KW	antiarteriosclerotic; anti-HIV; anti-allergic; cerebroprotective;		
Db	587	TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT	646	KW	antiparkinsonian; anticonvulsant; nootropic; neuroprotective;		
Qy	201	PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla	220	KW	antiinflammatory; ophthalmological; antithyroid; antiarthritic;		
Db	647	TTCATGCCCGAGACCCCGGCTTCTGCTGACTCAGCAGCAGCGCCAGGAGCCATGGCC	706	KW	antibacterial; viricide; protozoacide; antiparasitic; fungicide;		
Qy	221	AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGlu	240	KW	anorectic; cardiant; hypotensive; antiinfertility; hepatotropic;		
Db	707	GCCTCGGCTTCTGTGGGGCTCCGAGCAGGGCTGGGAGACCCGCCATCGGGCTGAG	766	KW	gene therapy; gene; ds.		
Qy	241	GihSerPheHileuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGly	260	XX	Homo sapiens.		
Db	767	CAGAGCTTTTCACTGGCCCTGCTGGCGGACCCCGGCACTACAGCCCTTCATCATCGC	826	OS	WO2003093427-A2.		
Qy	261	ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu	280	XX	13-NOV-2003.		
Db	827	GTCTCCCTGATGGCTTCCAGCAGCTGTGGGGGTCAACGCCGCTCATGTTCATGACAG	886	XX	30-APR-2003; 2003WO-US013629.		
Qy	281	ThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal	300	XX	30-APR-2002; 2002US-0376988P.		
Db	887	ACCATCTTTGAAGAGGCCAAGTTCAAGGACAGCAGCTGGCCCTGGTGTGTGGGTGTC	946	PR	14-JUN-2002; 2002US-0389095P.		
Qy	301	IleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArgLeu	320	PR	22-AUG-2002; 2002US-0405860P.		
Db	947	ATCCAGGTGTGTTCACAGCTGTGGCGGCTCTCATCATGACAGCAGCGGGAGGCTG	1006	PR	27-AUG-2002; 2002US-0406512P.		
Qy	321	LeuLeuValLeuSerGly-ValValMetValPheSerThrSerAlaPheGlyAlaTyrPh	340	XX	(INCY-) INCYTE CORP.		
Db	1007	CTCCTGGTCTTGTTCAGAGGTC-----	1028	PI	Swarnakar A, Tran UK, Khare R, Marquis JP, Hafalia AJA; Yue H;		
Qy	340	eLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSe	360	PI	Elliott VS, Becha SD, Ramkumar J, Kable AE, Wilson AD, Jin P, Jiang X, Jackson AA;		
Db	1028	-----	1028	PI	Forsythe IJ, Griffin JA, Bulloch SA, Jin P, Chang H, Richardson TW, Tang YT, Lee SY;		
Qy	360	rAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLe	380	PI	Mason PM, Chawla NK, Gietzen KJ, Fu GK;		
Db	1028	-----	1028	DR	WPI; 2004-022653/02.		
Qy	380	uPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle	400	DR	P-PSDB; ADH45412.		
Db	1028	-----	1028	XX	New human molecules for disease detection and treatment (MDDT) and		
Qy	400	ePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMe	420	XX	encoding polynucleotides, useful for diagnosing, preventing or treating		
Db	1028	-----	1028	XX	diseases associated with aberrant MDDT expression, e.g. cancer, stroke,		
Qy	420	tAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAl	440	XX	obesity or AIDS.		
Db				XX	Claim 5; SEQ ID NO 60; 267pp; English.		
Qy				XX	The invention relates to human molecules useful for disease detection and		
Db				CC	treatment (MDDT) and also the polynucleotides which encode and identify		
Qy				CC	and encode MDDT. The MDDT polypeptides and encoding polynucleotides have		
Db				CC	the following activities: cytostatic, antiarteriosclerotic, anti-HIV,		
Qy				CC	antiallergic, cerebroprotective, antiparkinsonian, anticonvulsant,		
Db				CC	nootropic, neuroprotective, antiinflammatory, antithyroid, viricide, protozoacide,		
Qy				CC	antiparasitic, fungicide, anorectic, cardiant, hypotensive,		
Db				CC	antiinfertility, and hepatotropic. The MDDT polynucleotide may be used in		
Qy				CC	gene therapy to treat disorders. The polypeptides and polynucleotides are		
Db				CC	useful in diagnosing, preventing or treating diseases or conditions		
Qy				CC	as autoimmune/inflammatory disorders (AIDS, allergies, atopic dermatitis,		
Db				CC	arthritis or thyroiditis), infections (e.g. bacterial, viral, parasitic,		

CC protozoal (or fungal), metabolic disorders (e.g. obesity), reproductive
 CC disorders (e.g. infertility), neurological disorders (Parkinson's
 CC disease, Alzheimer's disease, stroke, epilepsy), cardiovascular disorders
 CC (e.g. myocardial infarction or hypertension), eye disorders, or cell
 CC proliferative diseases (e.g. cancer, atherosclerosis, hepatitis). These
 CC are also useful in assessing the effects of exogenous compounds on the
 CC expression of nucleic acid and amino acid sequences of MDDT. The MDDT or
 CC its fragments are useful in screening compounds for effectiveness as
 CC agonist or antagonist of the polypeptides, or in altering the expression
 CC of the target polynucleotide and compounds that specifically bind to or
 CC modulate the activity of the polypeptide. The microarray of the invention
 CC is useful in monitoring or measuring protein-protein interactions, drug-
 CC target interactions, and gene expression profiles. This polynucleotide
 CC sequence represents the DNA of an MDDT polypeptide of the invention.
 XX
 SQ Sequence 1763 BP; 305 A; 543 C; 513 G; 402 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1.76e-142 Length: 1763
 Score: 1746.50 Matches: 353
 Percent Similarity: 74.00% Conservative: 0
 Best Local Similarity: 74.00% Mismatches: 1
 Query Match: 71.08% Indels: 124
 DB: 12 Gaps: 1

US-09-886-954A-1 (1-477) x ADH45466 (1-1763)

QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
 DB 63 ATGACCCCGAGACCCAGAGGAAACCCAGCGCCTTCCTGGGGCTCTCGCGGCGAG----- 118
 QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 DB 118 ----- 118
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaProPro 60
 DB 118 ----- 118
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValThrLeuGlyAla 80
 DB 118 ----- 118
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 DB 118 ----- 118
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 DB 118 ----- 118
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 DB 118 ----- 118
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 DB 119 -----GTCTACATCTCCGAATCGCTACCCAGCAGTCCGGGGTGTCTCGGCTCTCT 172
 QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 DB 173 GTGCAGCTAATGTGTCTGTGGCATCTCTGGCCCTACCTGGCAGCTGGGTGTCTGGAG 232
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 DB 233 TGGCGTGGCTGGCTGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 292
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 DB 293 TTCATGCCCGAGACCCCGGCTCTCTGTGCTGACTCAGCAGAGCGCCAGAGGCGCATGGCC 352
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheLeuAlaGlu 240
 DB 118 ----- 118

DB 353 GCCCTCGGCTTCTCTGGGGCTCTGAGCAGGGGTGGGAAGACCCCCCTCGGGGCTGAG 412
 QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIlePheProPheIleGly 260
 DB 413 CAGAGCTTTCACCTGGCCCTGTCTGGGAGCCCGGCATCTACAGGCCCTTCATCATCGGC 472
 QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 DB 473 GTCCTCCGTGATGGCTTCACGAGCTGTCTGGGGGTCAACGCCGTCTATGTTCTATGCGAG 532
 QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
 DB 533 ACCATCTTTTGAAGAGGCCAAGTTCAAGGACAGCAGCCTGGCCTCGTCTGGTGGTGTG 592
 QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgGluLeu 320
 DB 593 ATCCAGGTGCTGTTTACAGCTGTGGGGCTCTCATCATGACAGACAGGGCGGAGGCTG 652
 QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
 DB 653 CTCTCTGGTCTTGTCTCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 712
 QY 341 LysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
 DB 713 AAGCTGACCCAGGGTGGCCCTGGCAACTCTCTGCACGTGGCCATCTCTGGCGGCTGTCT 772
 QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCysLeu 380
 DB 773 GCACAGCTGT 832
 QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
 DB 833 TTCATCGCCGGCTTTTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGTGGGGT 892
 QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
 DB 893 TTCCCTCTGTCATGTCAAGGGCGTGGCGACAGGCATCTGCGTCTCTCATGTGCAGAGATC 952
 QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
 DB 953 GCCTTCTCTGTGACCAAGAGTTCAGCAGCTTCATGGAGTCTCTAGGCCCTATGAGGCC 1012
 QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 DB 1013 TTCTGGCTTGCCTCGCTTCTGTCATCTTCAGTGTCTTTTCACTTTTGTCTGTGTCCT 1072
 QY 461 GluThrLysGlyThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 DB 1073 GAACTAAAGGAAGACTCTGGAACAAATCACAGCCCATTTTGGAGGGCGGA 1123

RESULT 13

ABL90126

ID ABL90126 standard; cDNA; 1156 BP.

XX

AC ABL90126;

XX

DT 24-MAY-2002 (first entry)

XX

DE Human polynucleotide SEQ ID NO 688.

XX

KW Cytostatic; immunosuppressive; nootropic; neuroprotective; antiviral;

KW anti-allergic; hepatotropic; antidiabetic; antifungal; antitumor;

KW vulnary; anticonvulsant; antibacterial; antifungal; antiparasitic;

KW cardiant; gene therapy; cancer; immune disorder; cardiovascular disorder;

KW neurological disease; infection; human; secreted protein; gene; ss.

XX

OS Homo sapiens.

XX

XX WO200190304-A2.

XX

PD 29-NOV-2001.

XX

XX 18-MAY-2001; 2001WO-US016450.

XX 19-MAY-2000; 2000US-020551SP.
 XX (HUMA-) HUMAN GENOME SCI INC.
 XX Birse CE, Rosen CA;
 XX WPI; 2002-122018/16.
 DR P-PSDB; ABB89717.
 XX Novel 1405 isolated polypeptides, useful for diagnosis, treatment and
 PT prevention of neural, immune system, muscular, reproductive,
 PT gastrointestinal, pulmonary, cardiovascular, renal and proliferative
 PT disorders.
 XX Claim 4; SEQ ID NO 688; 2081pp + Sequence Listing; English.
 XX The invention relates to novel genes (ABL89449-ABL90853) and proteins
 CC (ABB89040-ABB90444) useful for preventing, treating or ameliorating
 CC medical conditions e.g. by protein or gene therapy. The genes are
 CC isolated from a range of human tissues disclosed in the specification.
 CC The nucleic acids, proteins, antibodies and (ant)agonists are useful in
 CC the diagnosis, treatment and prevention of: (a) cancer, e.g. breast and
 CC ovarian cancer and other cancers of the adrenal gland, bone, bone marrow,
 CC breast, gastrointestinal tract, liver, lung, or urogenital; (b) immune
 CC disorders e.g. Addison's disease, allergies, autoimmune haemolytic
 CC anaemia, autoimmune thyroiditis, diabetes mellitus, Crohn's disease,
 CC multiple sclerosis, rheumatoid arthritis and ulcerative colitis; (c)
 CC cardiovascular disorders such as myocardial ischaemias; (d) wound healing
 CC ; (e) neurological diseases e.g. cerebral anoxia and epilepsy; and (f)
 CC infectious diseases such as viral, bacterial, fungal and parasitic
 CC infections. Note: The sequence data for this patent did not form part of
 CC the printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 1156 BP; 177 A; 362 C; 345 G; 255 T; 0 U; 17 Other;

Alignment Scores:
 Pred. No.: 7,07e-137 Length: 1156
 Score: 1679.50 Matches: 344
 Percent Similarity: 72.12% Conservative: 0
 Best Local Similarity: 72.12% Mismatches: 10
 Query Match: 68.36% Indels: 125
 DB: 6 Gaps: 1

US-09-886-954A-1 (1-477) x ABL90126 (1-1156)

Qy 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlySerAla 20
 Db 53 ATGACGCCCGAGACCCAGGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGCGAR---- 108
 Qy 21 ProArgGlyArgValPheLeuAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
 Db 108 ----- 108
 Qy 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaProPro 60
 Db 108 ----- 108
 Qy 61 AlaProGluAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 108 ----- 108
 Qy 81 AlaAlaGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 Db 108 ----- 108
 Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 108 ----- 108
 Qy 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140

Db 108 ----- 108
 Qy 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 Db 109 -----GTCTACATCTCCGAATCCTACCCAGCAGTCCGGGGTTCCTGGCTCTCTGT 162
 Qy 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu 180
 Db 163 GTGAGCTAATGTGTCGTCGCACTCTCTGGCTTACTTGGCAGGCTGGTGGTCTGGAG 222
 Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 223 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGC 282
 Qy 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAla 220
 Db 283 TTCTATGCTCCAGACCCCGCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 342
 Qy 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
 Db 343 GCGCTGCGNTTCTGTGGGCTCCAGCAGGGCTGGGAGACCCCGCTGCTGGCTGGAG 402
 Qy 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIysProPheIleIleGly 260
 Db 403 CAGAGCTTTCACCTGGCCCTGCTGGCGCANCCTGGCCTCTACAGCGCTTTCATCATCG 462
 Qy 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 Db 463 GTCTCCCTGATGGCTTCCAGCAGCTGTGGGGGTCAACGCCGTCATGTTCTATGCAAG 522
 Qy 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValIleGlyVal 300
 Db 523 ACCATCTTGAAGAGGCCAGTTCAAGCAGCAGCCTGGCTGGCTGGCTGGCTGGCTG 582
 Qy 301 IleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
 Db 583 ATCCAGGTGCTGTTTACAGCTGTGGCGCTCTCATCTGCGCAGTGGCCTCTGGCGGCTG 642
 Qy 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
 Db 643 CTCTCTGGTCTTGTGAGGTGTGTGATGTTTTCAGCAGCAGTGTGCTTGGCGGCTACTTC 702
 Qy 341 LysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
 Db 703 AAGTGTCCAGGCTGGCGCTTGGCACTCTCTGCGCAGTGGCCTCTGGCGGCTGTCTCT 762
 Qy 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
 Db 763 GCACAGCCTGTGATGCCAGCGTGGGGCTNCGCTGGCTGGCGGCGAGCATGTGGCTTC 822
 Qy 381 PheIleAlaGlyPheAlaValGlyTyrGlyProIleProTrpLeuLeuMetSerGluIle 400
 Db 823 TTCTATGCGCGCTTGTGGCTGGCTGGCGGCGCCATCCCTGGCTCTCTCATGTCTGAGATC 882
 Qy 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
 Db 883 TTCCCTTGTTCATGTCAAGGGGTGGCGCAGCAGCAWYTCGCTCTCAMCAACTGGCTCATG 942
 Qy 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyVal 440
 Db 943 GSCTTTTCTYGGACCAAGAGTTTACAGAGCTTCATGGAGGCTCTCAGGCCCTATGAGGCC 1002
 Qy 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 Db 1003 TTCTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCTTGGCT 1061
 Qy 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db 1062 GAAACTAAAGGAAAGACTCTGGAAATAATCACAGCCCATTTTGGGGGCGCA 1112

RESULT 14

AAD12574

ID AAD12574 standard; cDNA; 1461 BP.

XX AAD12574;
XX
XX
XX 25-SEP-2001 (first entry)
XX
XX Human protein having hydrophobic domain encoding cDNA clone HP10784.
XX
XX Human; hydrophobic domain; gene therapy; nutritional supplement;
KW cell proliferation; immunomodulatory; autoimmune disorder; antimicrobial;
KW multiple sclerosis; rheumatoid arthritis; insulin-dependent diabetes;
KW haematopoiesis; tissue growth activity; Parkinson's disease; cytostatic;
KW Huntington's disease; Alzheimer's disease; chemotactic; chemokinetic;
KW haemostatic; thrombolytic; tumour growth inhibitor; anabolic;
KW contraceptive; antiinfertility; antiinflammatory; ss.
XX Homo sapiens.
XX
XX
XX Key Location/Qualifiers
FH CDS 61..849
FT /*tag= a
FT /product= "Human protein having hydrophobic domain"
FT /note= "CDS is specifically is claimed in claim 3"
XX
XX WO200149728-A2.
XX
XX 12-JUL-2001.
XX
XX 28-DEC-2000; 2000WO-JP009359.
XX
XX 06-JAN-2000; 2000JP-00000585.
XX 06-JAN-2000; 2000JP-00000588.
XX 11-JAN-2000; 2000JP-00002299.
XX 03-FEB-2000; 2000JP-00026862.
XX 03-MAR-2000; 2000JP-00058367.
XX
XX (PROT-) PROTEGENE INC.
XX (SAGA) SAGAMI CHEM RES CENT.
XX
XX Kato S, Kimura T;
XX WPI: 2001-418355/44.
XX P-FSDB; AAS06579.
XX
XX Human proteins with hydrophobic domains and the nucleic acids encoding
PT them, useful for preventing diagnosing and treating e.g. cancer,
PT Alzheimer's and inflammation.
XX
XX Claim 4; Page 287-289; 563pp; English.
XX
XX The present sequence is human protein with hydrophobic domain encoding
CC cDNA clone HP10784. The polynucleotide and polypeptide of the invention
CC may be used in the prevention, diagnosis and treatment of diseases
CC associated with inappropriate polypeptide expression. The polynucleotides
CC may be used to produce the polypeptide, by inserting the nucleic acids
CC into a host cell and culturing the cell to express the protein. The
CC polynucleotides and its complementary sequences may also be used as DNA
CC probes in diagnostic assays and also used in gene therapy. The
CC polypeptides may also be used as antigens in the production of antibodies
CC and in assays to identify modulators of polypeptide expression and
CC activity. The polypeptides and nucleic acids may be used as nutritional
CC supplements, to modulate cytokine and cell proliferation activity, to
CC modulate immune stimulation or suppression (e.g. for the treatment of
CC microbial infections and autoimmune disorders such as multiple sclerosis,
CC rheumatoid arthritis and insulin-dependent diabetes), to modulate
CC haematopoiesis, to modulate tissue growth activity (e.g. for the
CC treatment of Parkinson's disease, Huntington's disease and Alzheimer's
CC disease), to modulate activin and inhibin activity (e.g. for controlling
CC fertility), to modulate chemotactic and chemokinetic activity, to
CC modulate haemostatic and thrombolytic activity, to modulate receptor
CC ligand activity, to modulate inflammation and to inhibit tumour growth
XX
XX Sequence 1461 BP; 195 A; 501 C; 456 G; 309 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 7.84e-126 Length: 1461
Score: 1554.50 Matches: 324
Percent Similarity: 67.85% Conservative: 1
Best Local Similarity: 67.64% Mismatches: 1
Query Match: 63.27% Indels: 153
DB: Gaps: 2
US-09-886-954A-1 (1-477) x AAD12574 (1-1461)
QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
DB 61 ATGACGCCCGAGGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGCAGCGC 120
QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
DB 121 CCCCAGCGCGCGCGCTTCTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTC 180
QY 41 GlyPheAlaLeuGlyTyxSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
DB 181 GGCCTTCGGCTCGGCTACAGCTCCCGGCGCATCCCTAGCTGCAGCGCGCGCGCGCGCG 240
QY 61 AlaProArgLeuAspAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
DB 241 GCCCGCGCGCTGGAGCAGCGCGCGCTCTCTGTTTCGGGGCTGTCGTGACCTCTGGGTGCC 300
QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlySerLeu 100
DB 301 GCGCGCGGGGAGTGGTGGCGCGCTGGTGGTGGACCGCGCGCGCGCGCGCGCGCGCGCTC 360
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
DB 361 TTGCTGTGCTCGCTCGCTTTCGTCGGCGCTTTCGCGCTATCACCGCGCGCGCGCGCG 420
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
DB 421 TGGATGCTGCTGGGGGCGCGCTCTCTACCGCGCTGGCTGGCTGGGTGGTGGCTCTCTAGTG 480
QY 141 AlaProValTyxIleSerGluIleAlaTyxProAlaValArgGlyLeuLeuGlySerCys 160
DB 481 GCCCGGCTTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGGTTCGCGCTCTCTGT 540
QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyxLeuAlaGlyTrpValLeuGlu 180
DB 541 GTGAGCTAAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 600
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 601 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 660
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAla-MetAl 220
DB 661 TTCATGCCCGAGACCCCGCGCTTCTCTGCTGACTCAGCAGCGCGCGCGCGCGCGCGCTGC 718
QY 220 aAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGl 240
DB 718 ----- 718
QY 240 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyxLysProPheIleIleGl 260
DB 718 ----- 718
QY 260 yValSerLeuMetAlaPheGlnLeuSerGlyValAlaAlaValMetPheTyxAlaGl 280
DB 718 ----- 718
QY 280 uThrIlePheGluGluAlaAlaLysPheLysAspSerSerLeuAlaSerValValGlyVa 300
DB 718 ----- 718
QY 300 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgGlu 320
DB 718 ----- 718

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 25, 2005, 18:36:27 ; Search time 236 Seconds
(without alignments)
3307.220 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPEDPEETQPLGPPGSA.....CVPETKGTLEQTAHFEGR 477

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1202784 seqs, 818138359 residues

Total number of hits satisfying chosen parameters: 2405568

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+ p2n.model -DRV=xlp
-Q=/cgn2.1/USPTO spool_p/US09866954/runat_23022005_154243_13234/app_query.fasta_1.647
-DB=Issued Patents NA -QFMT=fastep -SUFFIX=p2n.rn1 -MINMATCH=0.1 -DOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09866954 @CGN 1.1 105 @runat_23022005_154243_13234 -NCPU=6 -ICPU=3
-NO WMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2.6/ptodata/1/ina/5A COMB.seq.*
2: /cgn2.6/ptodata/1/ina/5B COMB.seq.*
3: /cgn2.6/ptodata/1/ina/6A COMB.seq.*
4: /cgn2.6/ptodata/1/ina/6B COMB.seq.*
5: /cgn2.6/ptodata/1/ina/6C COMB.seq.*
6: /cgn2.6/ptodata/1/ina/6D COMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1215.5	49.5	740	4	US-09-949-016-3739
2	870.5	35.4	10143	4	US-09-949-016-15481
3	605	24.6	1545	4	US-09-489-039A-4731
4	578.5	23.5	1431	4	US-09-489-039A-4762
5	570.5	23.2	3915	4	US-09-023-655-1104
6	542.5	22.1	1752	4	US-09-679-686B-17
7	532.5	21.7	1675	4	US-09-679-686B-1
8	518	21.1	1960	4	US-09-679-686B-15
9	514.5	20.9	2856	3	US-09-643-597-135
10	514.5	20.9	2856	4	US-09-480-884A-135
11	514.5	20.9	2856	4	US-09-542-615A-135
12	514.5	20.9	2856	4	US-09-606-421B-135

ALIGNMENTS

RESULT 1
US-09-949-016-3739
; Sequence 3739, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 3739
; LENGTH: 740
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-3739

Alignment Scores:
Pred. No.: 3.29e-119 Length: 740
Score: 1215.50 Matches: 241
Percent Similarity: 99.59% Conservative: 0
Best Local Similarity: 99.59% Mismatches: 0
Query Match: 49.47% Indels: 1
DB: 4 Gaps: 1

US-09-886-954A-1 (1-477) x US-09-949-016-3739 (1-740)

QY 86 LeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuCySerVal 105

```

; ORGANISM: Human
; US-09-949-016-15481

Alignment Scores:
Pred. No.:      8,01e-81      Length:      10143
Score:          870.50      Matches:      429
Percent Similarity: 15.43%      Conservative: 0
Best Local Similarity: 15.43%      Mismatches: 3
Query Match:      35.43%      Indels:      2351
DB:              4          Gaps:      8

US-09-886-954A-1 (1-477) x US-09-949-016-15481 (1-10143)
QY      1 MetThrProGluAspProGluGlnProLeuLeuGlyProProGlyGlySer---- 19
Db      1291 ATGACGCCCGAGGACCCAGAGGAAACCCAGCGCTTCTGGGGCCTCTCTGGCGGAG-GTG 1349
QY      19 ----- 19
Db      1350 AGCTCCGGGGAAACCCGGGCGCGGATCGGCGCTCCCGCCCTCCCGCGTCCGCTCC 1409
QY      20 -----AlaProArgGlyArgValPheLeuAlaAla 30
Db      1410 GCTCACCTCGGCGCTGTCCTCCCGGCGCGCGCGCGCGCTCTTCTCGCGCGC 1469
QY      31 PheAlaAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyrSerSerProAla 50
Db      1470 TTGGCGCTGCTCCCTGGGCGCCACTCAGCTTCGGCTTCGGCTTACAGCTCCCGCGC 1529
QY      51 IleProSerLeuGlnArgAlaAlaProProAlaProArgLeuAspAlaAlaAlaSer 70
Db      1530 ATCCCTAGCTGCGCGCGCGCGCGCGCGCGCGCGCGCTGAGACACCGCGCGCTCC 1589
QY      71 TrpPheGly----- 73
Db      1590 TGGTTCGGGTGAGGCGCGCGCGCGCGCTCGCTCCCGCGCTGGGACCCAGCGCTCTCTC 1649
QY      73 ----- 73
Db      1650 GGGACGGGATCGGGACCTTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1709
QY      73 ----- 73
Db      1710 GCCGCCACACCTTTTCCCGACGAGACAGGATTTGGGCGCTCTCTGCCCCCATCCCTTCCC 1769
QY      73 ----- 73
Db      1770 TCGGGACGGGATCGAGCGCTTCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1829
QY      73 ----- 73
Db      1830 CTCAGCCCCCTACCCCTCCCTCGGGACGGGCGCGCGCGCGCGCGCGCGCGCGCGCGC 1889
QY      73 ----- 73
Db      1890 CCCGGGGAGTGGGTGGAGGGGAGGATGGGCTTCGGCGCGCGCGCGCGCGCGCGCGC 1949
QY      74 -----AlaValValThrLeuGlyAlaAlaAlaGlyGlyValLeuGlyGly 88
Db      1950 GCCGCCCGCTTCAGGCTGTGTGACCTTCGGGTGCGCGCGCGCGCGCGCGCGCGCGC 2009
QY      89 TrpLeuValAspArgAlaGlyArgGlySerLeuLeuLeuLeuLeuLeuLeuLeuLeu 108
Db      2010 TGGCTGGTGGACCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2069
QY      109 AlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeu 128
Db      2070 GCCGGCTTTGCGCTCATCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 2129
QY      129 LeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro----- 142
Db      2130 CTCACCGGCTGGCTTCGGGTGTGTGCTTCCCTAGTGGCGCGCGCGCGCGCGCGCGC 2189
;

1 CTGGGCGGCTGGTGGACCGCGCGCGCGCGCGCGCGCGCGCGCTCTTGGCTGTGCTCGGTG 60
106 ProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGly 125
61 CCCCTCGTGGCGGCTTTGGCGCTCATCACCGCGCGCGCGCGCGCGCGCGCGCGCTGCTGGG 120
126 GlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTyrIle 145
121 GGCCGCTCTCTCACCGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 180
146 SerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetVal 165
181 TCCGAAATCGCTACCCAGCAGTCCGGGGTGGCTCGGCTCTCTGTGTGACGCTAATGGTC 240
166 ValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGluTyrArgTyrLeuAla 185
241 GTCTGGGATCTCTCTGGGCTTACCTGGGAGGCTGGGCTGGGCTGGGCTGGGCTGGGCT 300
186 ValLeuGlyCysValProSerLeuMetLeuLeuLeuMetCysPheMetProGluThr 205
301 GTCTGGGCTGGTGGCGCGCGCGCGCGCGCGCGCGCGCGCTCTCTCATGTCTTCATGCCGAGACC 360
206 ProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeu 225
361 CCGCGCTTCTGTGACTCAGCACAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 420
226 TrpGlySerGluGlnGlyTyrGluAspProProIleGlyAlaGlu---GlnSerPheHis 244
421 TGGGGCTCCGAGCAGGCTGGGAGACCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGC 480
245 LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGlyValSerLeuMet 264
481 CTGGCGCTGTGGCGGAGCGCGCGCGCGCGCGCGCGCGCGCGCTTCATCGGCGTCTCCCTGATG 540
265 AlaPheGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePheGlu 284
541 GCCTTCAGCAGCTGTGGGGGTCAAGCGCGCTCATGTTCTATGACAGAGACCATCTTGAA 600
285 GluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyValIleGlnValLeu 304
601 GAGGCCAAGTTCAAGGACAGCAGCTGGCGTGGCGTGGCGTGGCGTGGCGTGGCGTGGC 660
305 PheThrAlaValAlaAlaIleMetAspArgAlaGlyArgArgLeuLeuValLeu 324
661 TTCACAGCTGTGGCGCTCTCATCATGACAGAGCGGCGGCGGCGGCGGCGGCGGCGGCT 720
325 SerGly 326
721 TCAGGT 726

RESULT 2
US-09-949-016-15481
; Sequence 15481, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CLO01307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15481
; LENGTH: 10143
; TYPE: DNA

```

Qy	142	-----	142
Db	2190	GAGTGTCTCTGTCTCGCGCCTGAGACCGAGGGGAGTGGGACAAACCGCTCCCCAGGCCT	2249
Qy	142	-----	142
Db	2250	GGGGCGCGGCTCCCCCTGGCGGGACCTTCTTGGGTGCCAGGTTGAAGTCCCTCGGTTAT	2309
Qy	142	-----	142
Db	2310	CTCGGGTCCCTCCGTCGACCTGGGAAGGATCCTACTGTCTCTCCATTTTACACTGA	2369
Qy	142	-----	142
Db	2370	GGTCATGACATCAGTCTCTGGAAAGGTGAAGTCTTTGCCCGAGGTCACAGCTAG	2429
Qy	142	-----	142
Db	2430	TCAGAGGGGAAGCAGTTGCAGGAACCCAGGTTGTCCACTTAGCCGTGCCCTCTTTTG	2489
Qy	142	-----	142
Db	2490	CTCTCAAACTGCGGATGATCCACAGAGCCACTCCCTACATTTTGGTTTTCATCCCT	2549
Qy	142	-----	142
Db	2550	GGCTTCGGGTCAATGACTGTCAATTAGCAGGAAGTTCCTGTCTCTGATTTCTCTGGGCAA	2609
Qy	142	-----	142
Db	2610	AGGCTGTGGCCTCTGTGTTCCCTTCTCACAGGCAGCCCTCTGGCTTCCCA	2669
Qy	142	-----	142
Db	2670	TCTTCCCCCTTCCCCCTCTGATCCAGGCTGCCACTTGGAGCCCGGAACATCTGGGT	2729
Qy	142	-----	142
Db	2730	CAAACTCTGTGTTCTGCCAATAACCTGACAGAGTGCCTCCCTGGGCTCAGTTTCTTCT	2789
Qy	142	-----	142
Db	2790	CTGAAAAACAGGGATTATGGCAATTAGCCCTGAGGTAGTGTCAAGATTTCAGTGGGACTT	2849
Qy	142	-----	142
Db	2850	TGGATAGACAGCCCTTGCATGGCTTGTAGAGGACTGATCTTCACAGCTTGGTGAACCA	2909
Qy	142	-----	142
Db	2910	TGAATGCCCCCAGCATGGGCTTGGACAAATATCTCCCCAAGTAGGTTGAAAGCAGCC	2969
Qy	142	-----	142
Db	2970	ACTCAGGCCCCCTCCCTGGGTGTCATTTACCAAAGTTGTAAACCGCAGCAGCAGCTG	3029
Qy	142	-----	142
Db	3030	GACCAGAAGGGTGGGAAGAGCTGGGCCCTTTTATGTCTTTGGTTTTTTTAACTGGGC	3089
Qy	142	-----	142
Db	3090	AGAAACACACCTGTGTTAGGGGCTTAAGCTATGGAATAACCTTAGATGCATGGGTGAGGG	3149
Qy	142	-----	142
Db	3150	CCTGTGGCCTGTGGGCCCAGGAGGAGCTAGTGGCAGCCACTGGTAGTGACCGAGGCAG	3209
Qy	142	-----	142
Db	3210	AGATTCTGTGTTCTAGTCCAAAGCCCACTCAAGCGTGGGCCCTTGGGCTAGGATCTCT	3269
Qy	142	-----	142

Db	3270	GGGTGGCGTCA	CAGTGGAGCTGGTTAA	CAAGGTGGAATCTTT	CGCCTGTCCAGACCTPACT	3329	
Qy	142	-----	-----	-----	-----	142	
Db	3330	GCATCGGAATCT	ATGATTTATCCAGGGTGTCT	GTGCACATTAACGAAGGCTT	GAGAAGC	3389	
Qy	142	-----	-----	-----	-----	142	
Db	3390	TCTGTGCCACCG	CGCTCTCGTTCATATATG	CAGCCCTCTCGTTACCTTTA	ATCCCCCTTT	3449	
Qy	142	-----	-----	-----	-----	142	
Db	3450	AAACACCTGCCA	CAGATCTTTACGTGTGGGAC	CTGGGATGGCTCTGCTCACT	ACCTAGGC	3509	
Qy	142	-----	-----	-----	-----	142	
Db	3510	AGGTTCCNCTA	AGGGGCTTCAGGSCCTG	CCGTAAACCCCTTGC	AAAGGACTTTCCTCAGCTC	3569	
Qy	142	-----	-----	-----	-----	142	
Db	3570	TGAATGGCTG	AGACCTTTCCATTTTAAGA	ACTCTTGGAGCTGTGATAAG	CGCTCATCTTT	3629	
Qy	142	-----	-----	-----	-----	142	
Db	3630	TTTTTTTTTTTT	TAATTTTATTTAA	AAATTTTCGAGTTTCG	CTCTTTGTGCCAGGCT	3689	
Qy	142	-----	-----	-----	-----	142	
Db	3690	GGAGTGCAAT	TGGTGGATCTTTGGCTC	ACCGAACCTCCACTCG	AGGGTTCAAGTGATTC	3749	
Qy	142	-----	-----	-----	-----	142	
Db	3750	TCCTCCCTCA	ACCTCCTGAGTAGTGGG	TATACAGATGTGTGCC	ACTAGGCCAGCTAAT	3809	
Qy	142	-----	-----	-----	-----	142	
Db	3810	TTTGTATTTT	PAGTAGATGGGTTTCT	CCCTGTTGGTCAGGCC	AGTCTCAAACTCCCA	3869	
Qy	142	-----	-----	-----	-----	142	
Db	3870	ACCTCTGGT	GTATCTGCCCGCTCG	CGCTCCACAGAGTCT	GGGATTTGCAGGCATG	AGCCAC	3929
Qy	143	-----	-----	-----	-----	150	
Db	3930	TGCGCCAGCA	TAAATCCTCATCTG	ATTGCTGGCAGTCTCA	CTCTCCGAATCGCTA	3989	
Qy	150	rProAlaVal	ArgGlyLeuLeuGlySer	CysValGlnLeuMetVal	ValValGlyIleIe	170	
Db	3990	CCAGCAGTCC	GGGGTTGCTCGCTCCT	GTGTGCAGCTAATGGT	CGTCGTGGCATCCT	4049	
Qy	170	uLeuAlaTyrIeu	-----	-----	-----	174	
Db	4050	CCTGGCTAC	CTGGCAGGTATAGTT	GCATTATCTTGCTT	CTGTTGTGTGCATGTG	4109	
Qy	174	-----	-----	-----	-----	174	
Db	4110	GCCATTTT	CAGATGTGTAAACAG	AGGTTTCAGTGAGGT	CAAGGCACTTGTCTC	AGGTCAC	4169
Qy	174	-----	-----	-----	-----	174	
Db	4170	ATAGCCAGT	GCCCAACATGCCAG	CTCGCTTCTGGGAC	CTTTGGGATAGGT	CAGTCTTT	4229
Qy	174	-----	-----	-----	-----	174	
Db	4230	TTTTTTTTTT	TTTGACACAGATCC	CGCTCTGTGCGCC	AGGCTGGAGTGC	AGTGGTGC	4289
Qy	174	-----	-----	-----	-----	174	
Db	4290	TCTCTGCA	ACCTCTGCCCTCCT	GGGTTCAAGCGA	TTCTCTGTCTC	AGCTTCCCAAGTAGC	4349
Qy	174	-----	-----	-----	-----	174	

Db 4350 TGGGACTACAGGCATGTGCCACCACACCTGGCTAAATGTTTGTATATTTTGGTAGAAGGG 4409
Qy 174 ----- 174
Db 4410 GGTTCACCGTGTAGCCAGGATGGTCTTGATCTCCTGACCTCATGAACCCCAACCTC 4469
Qy 174 ----- 174
Db 4470 AGCTCCCAAGTGTGGGATTACAGGTTGAGCCACCGCCCGCGATAGGTGTGC 4529
Qy 174 ----- 174
Db 4530 TTATGACTGAGTTCAGGGACAGCTGGAGAAGGTATACCTCAGCTTTCCTGAAGCCCC 4589
Qy 174 ----- 174
Db 4590 CTCTAGGGGTGGGAAGGGCTGACTGCTTAGGTCACTTACTGTGAGTGCACAAAC 4649
Qy 174 ----- 174
Db 4650 CAGAGGGCTCCAGAGTGGGTGGTCCCTGTAAATAATCTTCAGCCCGGGCGAGTAAAGGG 4709
Qy 174 ----- 174
Db 4710 CTGAGCAAGTCCCTGAGAAATGTAGTGCAGATTCAAGATGTGGCGAGTTGCACGTG 4769
Qy 174 ----- 174
Db 4770 AAGGAGCCGGGTCTGGGACCCCGAGGGCAAAAGCTGGGCCAGCATGAAGAACTTC 4829
Qy 174 ----- 174
Db 4830 AGGAGACAGGTGTGTTCTCATTTAGAAGAGCCAGGCTTTAGAGCCAGACCCAGTGCGT 4889
Qy 174 ----- 174
Db 4890 CACACCTGTAAATCCAGCACTTTGGGAGGCAGAGGTGGCGGATCACITGAGGTCAGGAG 4949
Qy 174 ----- 174
Db 4950 TTGAGATCAGCCTGGCAACATGTGTAAGAACCCATTTCTATTTAAATFACAAAATTAG 5009
Qy 174 ----- 174
Db 5010 CCAGGCGTGTGTGGGCACCTGTAATCCAGCTACTGGGGGCTGAGGCACAGAAATT 5069
Qy 174 ----- 174
Db 5070 GCTTGAACCCAGAGGCTGAGGTTTCAGTGAGCTGAGATCGGCCACTGCACCTCCAGTGT 5129
Qy 174 ----- 174
Db 5130 GGGCGCACAGTGAAGTCTGTCTCAAAAGAAAGAAAAAACAAGGCGAGGCTTCAATC 5189
Qy 174 ----- 174
Db 5190 TGACCATCTGAGTTGGGATAGCTCTCTAGCTGCTGTGTGACCTTGGCCAAAGTCAT 5249
Qy 174 ----- 174
Db 5250 ACCACCTCTGAAGCCTCAGTTTCTCTGTCTTGAATGGGCATAACCTCATCTAGTCCAG 5309
Qy 174 ----- 174
Db 5310 TACCTGTCATTAAATTAATTGACATGAAGAGTGCCTAACAGTGGGTTCCAAATGAGA 5369
Qy 174 ----- 174
Db 5370 ATTGTCAAGGATGACTTGAGAACCATGGGATCATCGGGGTAGAGGCCCTCGGGGCCCT 5429
Qy 174 ----- 174
Db 5430 GACGGGCTATGGCTGGCCCGAGGCAACAGCCCTCTGCATCTGTCAGGGCGGATGATACAC 5489

Qy 174 ----- 174
Db 5490 AGGTCTTTGGTGTTCATTTTACCACCTTAAACCCTTAAACAAATTCAGAAATCTCGCCA 5549
Qy 174 ----- 174
Db 5550 GCGATCTGGCTACTTTGGGAGACCTGAGTAACCAAGCCCCCAGCTGTGTGCGAGGTGTG 5609
Qy 174 ----- 174
Db 5610 ACATTCATCATTTTCGGGACAGGTACTGGGTGTCCACCACCGTGGCACTATCTTCAGC 5669
Qy 174 ----- 174
Db 5670 TCATCTTAATATGTGTGTCT 5729
Qy 174 ----- 174
Db 5730 ATCCACTGGGTGGTGAACCTCTGATAAATCCACAAAGTGACATGGTGTGGCACGCTTG 5789
Qy 174 ----- 174
Db 5790 CTTACATTTGGGCAATCATGTTTCCAAAGACTAGACTTCATAGAACAAATGTCTGCACCC 5849
Qy 174 ----- 174
Db 5850 CCAGCAGAGAAGAAAATAGGGACATCATGGCTACCTTAGAGGTATTTTTCCACTGGAA 5909
Qy 174 ----- 174
Db 5910 CCAGGATGGTAAGCCTTGGGCTTAGGAAAAATTAGTAATTAGAGCAGCAATCTCCTAG 5969
Qy 174 ----- 174
Db 5970 ACTCATTTATCCAGGGCCCCCTTTGTAAGAGAACCTTTTAACCTTATATAAACTAGTA 6029
Qy 174 ----- 174
Db 6030 ATTCTGCACCTTATGCTGTCTTTCATAGCTCAAGACAGCAAAATCCTCTCCTCTCTACT 6089
Qy 174 ----- 174
Db 6090 TTCTATGATACTTTGAATAGATAATGCAATCAAAATGACTCAACACTCAAGTGCAGAGTG 6149
Qy 174 ----- 174
Db 6150 GAGAGTGTCCAGTGAAGGCTTCTCCGCTCTGTCCCTTTCTTGTCTGGGGAGTCCC 6209
Qy 174 ----- 174
Db 6210 CATTCCTGTTTCTTGTGTAGCTTTCAGAGAGTCTGTTTATATAGGGGACATGCATATTTA 6269
Qy 174 ----- 174
Db 6270 TCCTCCCTCCCCACCTCTTCTCAAAATGGTGGCCACATCATACCATGGGACACGACGA 6329
Qy 174 ----- 174
Db 6330 AGCACCTCATCTTCATTTTGTGGTGTACATAAATGTTTCTGTGAAACGATGCAGCTAG 6389
Qy 174 ----- 174
Db 6390 TAAATGTCGGGTGTGTATACCACTGTGCTTCGCGAGCTGGTGTGATTTGGCCCTGCCA 6449
Qy 174 ----- 174
Db 6450 TGTGCTTCATGAACAGTTGTGATTAATGCAATGTGCAATGTGAGCAACATGAGGCGAGATGT 6509
Qy 174 ----- 174
Db 6510 GCAGAGGCGAGCGAGGCCCATCTCTGAAGCTGAGTCTTTTGTAGGCTGGCAAGGCCAGGCT 6569


```

; LENGTH: 1545
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
; US-09-489-039A-4731

Alignment Scores:
Pred. No.: 5,47e-54 Length: 1545
Score: 605.00 Matches: 156
Percent Similarity: 51.60% Conservative: 86
Best Local Similarity: 33.26% Mismatches: 187
Query Match: 24.62% Indels: 40
DB: 4 Gaps: 9

US-09-886-954A-1 (1-477) x US-09-489-039A-4731 (1-1545)

QY 20 AlaProArgGlyArgArg-----ValPheLeuAlaAlaPheAlaAlaAla 34
Db 154 TCGCCGCGGACGCAACGCTGATACCCGCGGATGAACTGGTTGTTTCCATCGCGCGGCG 213
QY 35 LeuGlyProLeuSerPheGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeu 54
Db 214 GTGCTGGCTTGTCTTTGGCTGGATATCGGGGTGATATCCGGAGGTTGGCCCTTTATA 273
QY 55 GlnArgAlaAlaProProAlaProArgLeuAspAlaAlaAlaAlaSerTrpPheGlyAla 74
Db 274 ACCGACCATTTTCCCTTATCCAGCCAGCTTCAGGAG-----TGGGTGGTTAGC 321
QY 75 ValValThrLeuGlyAlaAlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAla 94
Db 322 AGCATGATGTGGGGGGCGGATAGGCGCGCTGTAAACGGCTGGCTGTCTTTCCGCGCTT 381
QY 95 GlyArgLysLeuSerLeuLeuLeuCysSerValProPheValAlaGlyPheAlaValIle 114
Db 382 GCGCGTAATACAGCTGATGGCGGGCGGTACTCTTTGTCGCGGCTCTATCGGATCC 441
QY 115 ThrAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCys 134
Db 442 GCTTTTGGCGCCAGCGTGGAGGTGCTGCTGGTGGCGCGGTGGTGGTGGCGGAGTC 501
QY 135 GlyValAlaSerLeuValAlaProValTyrIleSerGluIleAlaIleAlaValArg 154
Db 502 GGGATTGGCTTTATACCGCGCGCTGTACTCTCCGAGATGGCCAGGAGACGTCGCG 561
QY 155 GlyLeuLeuGlySerCysValGlnLeuMetValValGlyIleLeuLeuAlaIleVal 174
Db 562 GGGAAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 621
QY 175 AlaGlyTrpValLeuGluTrp-----ArgTrpLeuAlaValLeuGly-----CysVal 190
Db 622 TCCGATACCGCTTTAGTACGCGGTAACTGGCGCGCCATGCTGGCGGTGCTGGCGCTG 681
QY 191 ProProSerLeuMetLeuLeuMetCysPheMetProGluThrProArgPheLeuLeu 210
Db 682 CGGCGGTGATCTGATCTGCTGCTCTTTTGGCAACAGCCGCGCTGGCGTGGCG 741
QY 211 ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrpGlySerGluGln 230
Db 742 GAGAGGAGCAGCATATATCAAGAGAGAGAGTGTGGGATGTGCTGGCGATACCTCGGA 801
QY 231 GlyTrpGluAspProProIleGlyAlaGluGlnSerPheHisLeu----- 245
Db 802 AAGCGCGCGGACGAGCTTAAACGAGATCGTGGAGGCTCAAGCTGAAGCGGCGGTGG 861
QY 246 AlaLeuLeuArg---GlnProGlyIleTyrLysProPheIleIleGlyValSerLeuMet 264
Db 862 GGGTGTGTTAAGATCAATCGTAACGTCGCGCGCGGTGTTCTCTGGGATGCTGCTGCG 921
QY 265 AlaPheGlnGlnLeuSerGlyValAlaAlaValMetPheTrpAlaGluThrIlePheGlu 284
Db 922 GCGATGCGAGCAATTCCCGGCGATGAACATCATGATGATGATGATGATGATGATG 981
QY 285 GluAlaLysPhe-----LysAspSerSerLeuAlaSerValValGlyValIle 301

```

```

Db 8729 GTCAGGCTTCTGCGGCCAAGGTGCCAGAACACCAAGCTGCGCCGTTCTCTGGAGCGCGG 8788
QY 325 ----- 325
Db 8789 CTTGCCCTCTCAGAGGATCCAGGCCATGCTGTCGCGCCCTGGGCGCTTCCCGGGTGC 8848
QY 326 -----GlyValValMetValPheSerThrSerAlaPh 336
Db 8849 CCGGCCACTGCGCGCCCTCCGCGCTGCAGGTGTGCTCATGTGTTCAGCAGAGTGCCTT 8908
QY 336 eGlyAlaTyrPheTyrLeuThrGlnGlyProGlyValMetSerSerHisValAlaIleSe 356
Db 8909 CCGCGCCCTACTTCAAGCTGACCCAGGTGGCCCTGGCACTTCTCGACGTGGCACTTC 8968
QY 356 rAlaProValSerAlaGlnProValAlaSerValGlyLeuAlaTrpLeuAlaValGln 376
Db 8969 GCGGCTGTCTCTGCACAGCTGTGATGATGATGATGATGATGATGATGATGATGATG 9028
QY 376 ySerMetCysLeuPheIleAla----- 383
Db 9029 CAGCATGTGCTTCTTCATGCG- CGGTAAAGGGGCGCTGTGGAGGCTGGCGGAGGTGG 9087
QY 383 ----- 383
Db 9088 AGGTGATCCCCCGGCCCTGTGTGGCGGCTCCGCGAGTGGGTCTTCCCGACCTCA 9147
QY 383 ----- 383
Db 9148 CCTCTCCTCTGTAAGTTCCAAGCTGTGATTTCTAGCTCTGACCATGAGCTGTGGA 9207
QY 383 ----- 383
Db 9208 ACTAGGCCCTCTGCGTCTCTTCTTCACTTAATCCCTGAGCAGCCCTCGAGCAGGCCCT 9267
QY 383 ----- 383
Db 9268 TGTGACCCCATCTCAGACAGAGCGCTGTGCTGAGGAGGAGAGGCTGGCGAGGCT 9327
QY 383 ----- 383
Db 9328 TGGCAGAAAGCATGGAGCTGTGGGACCACTAGGATCCGGGACCCCAAGCCCA 9387
QY 383 ----- 383
Db 9388 GGTGGGAGCTGGTCTCTGCTCTGCGACAGCTGAGAGCCCGCGGCTGGGAGCC 9447
QY 384 -----GlyPheAlaValGlyTrpGlyProIlePr 393
Db 9448 TGATGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 9507
QY 393 oTrpLeuLeuMetSerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCy 413
Db 9508 CTGCTCTCATGTGAGATCTTCCCTCTGCAATGTCAGAGGCTGGCGGAGGAGGATCTG 9567
QY 413 sValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMet 432
Db 9568 CGTCTCACCACACTGGCTCATGCGCTTCTGTCGACCAAGAGTTTCAGCAGCGCTCATG 9625

```

RESULT 3
 US-09-489-039A-4731
 ; Sequence 4731, Application US/09489039A
 ; Patent No. 6610836
 ; GENERAL INFORMATION:
 ; APPLICANT: Gary Breton et. al
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
 ; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
 ; FILE REFERENCE: 2709.2004001
 ; CURRENT APPLICATION NUMBER: US/09/489,039A
 ; CURRENT FILING DATE: 2000-01-27
 ; PRIOR APPLICATION NUMBER: US 60/117,747
 ; PRIOR FILING DATE: 1999-01-29
 ; NUMBER OF SEQ ID NOS: 14342
 ; SEQ ID NO 4731

```
Db 982 ATGGCGGGCTTTACCACTACTGAAACAGCAGATGATCGCACCTCTGGTGGCGCTGAC 1041
Qy 302 GlnValLeuPheThrAlaValAlaLeuLeuMetAspArgAlaGlyArgLeuLeu 321
Db 1042 TTTATGTTTGGCCACCTTTATGGGGTTCACAGTGTGTAAGCGGCGCGAAGCCAGCG 1101
Qy 322 LeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLys 341
Db 1102 CTGAAATCGGCTTTAGCGGTGATGGCGCTGGGCACTCTGGTGTGGGC-----TACTGC 1155
Qy 342 LeuThrGlnGlyProGlyAenSerSerHisValAlaLeuSerAlaProValSerAla 361
Db 1156 CTGATCATGTCACATGCGCACC-----1179
Qy 362 GlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValAlaGlySerMetCysLeuPhe 381
Db 1180 -----GCATCCAGCGGCTCTCTGGCTCTCGTGGCATGACCATGATGTG 1227
Qy 382 IleAlaGlyPheAlaValGlyTyrGlyProIleProTyrLeuLeuMetSerGluLeuPhe 401
Db 1228 ATTGGCGGTATGCGATGAGCGCGCGGTGGTGTGGATCTCTGCTCCGAGATCCAG 1287
Qy 402 ProLeuHisValGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAla 421
Db 1288 CCGCTAAATGCGCGACTTCGGTATCATCTGCTCGACCCACCACTGGGTGTCGAAC 1347
Qy 422 PheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPhe 441
Db 1348 ATGATCATCGGCGCACTTCTGAGCGTGTGTGCGGATGGCGCGCGCCGACCTTC 1407
Qy 442 TrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGlu 461
Db 1408 TGCTCTACACGCGCTCAACGTGGCTTTATCGGCATCATCTCTGGCTGATCCGGAA 1467
Qy 462 ThrLysGlyLysThrLeuGluGlnIle 470
Db 1468 ACCAAGATGTCACTCCCTCGACACATT 1494

RESULT 4
US-09-489-039A-4762
; Sequence 4762, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ. ID NOS: 14342
; SEQ. ID NO 4762
; LENGTH: 1431
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-4762

Alignment Scores:
Pred. No.: 3,11e-51 Length: 1431
Score: 578.50 Matches: 149
Percent Similarity: 48.74% Conservative: 84
Best Local Similarity: 31.17% Mismatches: 180
Query Match: 23.54% Indels: 65
DB: 4 Gaps: 10

US-09-886-954A-1 (1-477) x US-09-489-039A-4762 (1-1431)

Qy 27 PheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyr 46
Db 82 TTCTCGTCTGTTCTCCCGCGCTGGCTGGCTGCTCTGCTGCTGATCGGTGTT 141
Qy 47 SerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaProAlaProAlaAsp 66
```

```
Db 142 ATGGCGGGTGTACCTTTATTTGCAATGAGTTC-----CAGATTTCCGCG 189
Qy 67 AlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyValAlaAlaGlyGlyValLeu 86
Db 190 CACACCCAGAGTGGGTGTCAGCTCCATGATGTTCCGGGCTGCGCGCGGGTCCG 249
Qy 87 GlyGlyTyrLeuValAspArgAlaGlyArgLysLeuSerLeuLeuCysSerValPro 106
Db 250 ACGCGTGGCTCTCTTTCAAACTGGCGCGGAAAGAGCTGATGATCGCGCCATCCTC 309
Qy 107 PheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGly 126
Db 310 TTGCTCGCGGTGCTGCTCTCTCGCGCGCGCAACAGCTCGAGATCTCTGCTGGTTCC 369
Qy 127 ArgLeuLeuThrGlyLeuAlaCysGlyValAlaAlaSerLeuValAlaProValTyrIleSer 146
Db 370 CGTGTGCTGCTCGGCTGGCGGTGGCTCATATACGGCTCCGCTGATCTGTGTCG 429
Qy 147 GluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetValVal 166
Db 430 GAAATCGCCCCAGAAAAAATTCGCGCAGTATGATTTCCATGATACAGCTGATGATCACC 489
Qy 167 ValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGlu-----TrpArg 182
Db 490 ATCGGGATCTTGGCGCTATCTCTGTACACCGCTTTCAGCTACAGCGCGCATGGCG 549
Qy 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
Db 550 TGGATGCTCGGGGTATCATCATTCGGCGGTGTTGCTGCTGATCGCGCTTATCTTCC 609
Qy 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaLeu 222
Db 610 CCGGACAGCGCGCTGGTTCGCCGCAACCTGCTTGTGTCGATCGGAAACGCTGCTG 669
Qy 223 -----ArgPhe 224
Db 670 CTGCGCTGCGGATACCAAGCGGAAACGCGAGCTCGATGAAATCCGTGAAAGC 729
Qy 225 LeuTrpGlySerGluGlnGlyTyrGluAspProProIleGlyAlaGluGlnSerPheHis 244
Db 730 CTGAAGGTAAACACAGTCCGGCTG-----753
Qy 245 LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIle-----IleGlyValSerLeu 263
Db 754 ---TCGCTGTTAAAGACACACAGCACTTCGCCCGCGGTGTTCTCGGCATCTGCTG 810
Qy 264 MetAlaPheGlnGlnLeuSerGlyValAenAlaValMetPheTyrAlaGluThrIlePhe 283
Db 811 CAGGTGATGCAACAGTTCACCGGATGAACGTCATCATGTACTACGCGCGGAAGATCTTT 870
Qy 284 GluGluAlaLysPheLysAspSerSer-----LeuAlaSerValValGlyVal 300
Db 871 GAGTGGCGGGTTATGCCAACACCACTGAGCAAAATGTGGGGGACACAGTGTGCGGTCTG 930
Qy 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLeu 320
Db 931 ACTACGTGCTGGCCACCTTTATCCCATCGTCTGGTTCGACCGCTGGCGCGGTAACCG 990
Qy 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
Db 991 ACGCTGATCTTGGCTTTATCTGTATG-----GCCGCGGGA-----1026
Qy 341 LysLeuThrGlnGlyProGlyAenSerSerHisValAlaIleSerAlaProValSer 360
Db 1027 -----ATGGCGCTCTGGGTACCATGATGACATCGGCATTCACCTCTACCCG 1077
Qy 361 AlaGlnProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCysLeu 380
Db 1078 -----CAGTACATCGCGCTCTGATGCTGCTGATG 1107
Qy 381 PheIleAlaGlyPheAlaValGlyTyrGlyProIleProTrpLeuLeuMetSerGluIle 400
Db 1107 -----CAGTACATCGCGCTCTGATGCTGCTGATG 1107
```


Db 1108 TTCATGTCGGTTGCGCATGAGCGCGCCCACTGATTGGGTACTGTGCTCCGAATC 1167
Qy 401 PheProLeuHisValLysGlyValAlaThrGlyLeCysValLeuThrAsnTrpLeuMet 420
Db 1168 CAGCGCTGAAAGCGCGCACTTCGGTATACCTGCTCCACAGCAGCAACTGATGATGCC 1227
Qy 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
Db 1228 AACATGATGTCGGCGCCACCTTCCTGACCATGCTCACTGCTGGCAGCGCCCAATACC 1287
Qy 441 PheTrpLeuAlaSerAlaPheCysLePheSerValLeuPheThrLeuPheCysValPro 460
Db 1288 TTCCTGGGTGACGGCGGTCTGAACGTGCTGTTATCTCTGCTGACGCTGTGGCTGATCCCG 1347
Qy 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHis---PheGluGlyArg 477
Db 1348 GAAACCAAAACGTCTCGCTGGAACATATTGAACGTAAACCTGATGACGGGTCTGT 1401

RESULT 5

US-09-023-655-1104
; Sequence 1104, Application US/09023655
; Patent No. 6607879
; GENERAL INFORMATION:
; APPLICANT: Cocks, Benjamin G.
; APPLICANT: Susan G. Stuart
; APPLICANT: Jeffrey J. Seilhamer
; TITLE OF INVENTION: COMPOSITION FOR THE DETECTION OF BLOOD CELL GENE
; TITLE OF INVENTION: EXPRESSION
; NUMBER OF SEQUENCES: 1508
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: INCYTE PHARMACEUTICALS, INC.
; STREET: 3174 PORTER DRIVE
; CITY: PALO ALTO
; STATE: CALIFORNIA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Word Perfect 6.1 for Windows/MS-DOS 6.2
; CURRENT APPLICATION DATA: US/09/023,655
; FILING DATE: HEREWITH
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: Zeller, Karen J.
; REGISTRATION NUMBER: 37,071
; REFERENCE/DOCKET NUMBER: PA-0001 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (650) 855-0555
; TELEFAX: (650) 845-4166
; INFORMATION FOR SEQ ID NO: 1104:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3915 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: GENBANK
; CLONE: g183684
US-09-023-655-1104

Alignment Scores: 1.1e-49 Length: 3915
Pred. No.: 570.50 Matches: 152
Score: 49.90% Conservative: 88
Best Local Similarity: 31.60% Mismatches: 190
Query Match: 23.22% Indels: 51

DB: 4 Gaps: 9
US-09-886-954A-1 (1-477) x US-09-023-655-1104 (1-3915)
Qy 26 ValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGly 45
Db 270 CTGATATTGGCCATCACAGTTGCTACATCGGCTCTTCCAAATTTGGCTACCAACTGGG 329
Qy 46 TyrSerSerProAlaIleProSerLeuGlnArg----- 56
Db 330 GTCATCAATGCTCCTCAGAGAAGATCAATAAGAAATTTATCAATAAACTTTGACGACCAAG 389
Qy 57 ---AlaAlaProAlaProArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaVal 75
Db 390 GGAATATGCCACCCCTCTGAGGTGCTGCTACAGTCTCTCTGCTCTCTGCTGTCGTCATA 449
Qy 76 ValThrLeuGlyAlaAlaAlaGlyValLeuGlyTrpLeuValAspArgAlaGly 95
Db 450 TTTTCGTCGGGGGTATGATCGGCTCTTTTCGTCGGAATCTCTGCTCAACGCTTTGGC 509
Qy 96 ArgLysLeuSerLeuLeuLeuCysSerValProPheValAlaGly-----PheAla 112
Db 510 AGGCGCAATTCAATGCTGATTTGCAACCTGTGGGTGCTCACTGGTGGCTGCTTTATGGGA 569
Qy 113 ValIleThrAlaAlaGlnAspValTrpMetLeuGlyGlyArgLeuLeuThrGlyLeu 132
Db 570 CTGTGTAAAGTAGCTAAGTCGGTTGAAATGCTGATCTCGGTGCTGCTGTTATTTGGCCTC 629
Qy 133 AlaCysGlyValAlaSerLeuValAlaProValTyrIleSerGluIleAlaTyrProAla 152
Db 630 TTCTGCGGACTCTGCACAGGTTTGTGCCCATGATCATATTGGAGAGATCTCGCCTACTGCC 689
Qy 153 ValArgGlyLeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAla 172
Db 690 CTGCGGGTGCCCTTTGGCACTCTCAACCACTGGGATGCTTCTTGAATTTCTGGTGGCC 749
Qy 173 TyrLeuAlaGly-----TrpValLeu-----GluTrpArgTrpLeuAlaValLeuGly 188
Db 750 CAGATCTTTGCTGGAATTCATCTTTGGGTCTGAAAGAGCTATGGCGGCTGCTACTGGGT 809
Qy 189 -----CysValProProSerLeuMetLeuLeuMetCysPheMetProGluThrPro 206
Db 810 TTTTACCATCTCTCTGCTATCTTACAAAGTGCAGCCCTTCCATTTTCCCTGAAAGTCCC 869
Qy 207 ArgPheLeuLeuThrGlnHisArgArgGlnGlu---AlaMetAlaAlaLeuArgPheLeu 225
Db 870 AGATTTTGTCTATTACAGAAAAGAGAGAGAATGCTAAGCAGATCTCCACGCGGTG 929
Qy 226 TrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGluGlnSerPheHis--- 244
Db 930 TGGGGCACCCAGGATGATATCCCAAGACATCCAGAGATGAAAGATGAGATGCAAGGATG 989
Qy 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrLysPro 256
Db 990 TCACAAAGAAAGCAAGTCAACGCTGCTAGAGCTCTTTAGAGTGTCCAGTACCCGACAGCC 1049
Qy 257 PheIleIleGlyValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaValMet 276
Db 1050 ATCATCAATTTCCATTTGCTGCTCAGCTCTCTCAGCAGCTCTCTGGGATCAATGCTGTGTC 1109
Qy 277 PheTyrAlaGluThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerVal 296
Db 1110 TATTACTCAACAGGAATCTTCAAGATGAGGTGTTCAAGAGCCCATCTATGCCACCATC 1169
Qy 297 ValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAla 316
Db 1170 GCGCGGGTGTGTTAATAACTATCTTCACTGTAGTTTCTCTATTCTCGTGGAAAGGCA 1229
Qy 317 GlyArgArgLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPhe 336
Db 1230 GGAAGAAGGACTCTGCATATGATA----- 1253
Qy 337 GlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSer 356

QY 376 GlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeu 395
Db 1239 CTCCTCATGCTCTATGTGGCGGGTTCGTGGTGGTGGCGCCCTGGGGTGGCTG 1298
QY 396 LeuMetSerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeu 415
Db 1299 GTGCCAGCGAGATCTCCCGCTGGAGATCAGCGCGCGGGGAGAGATCAAGCTGTGCG 1358
QY 416 ThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeu 435
Db 1359 GTACATGCTCTTACCTTGTCTATCGCGAGGGTTCCTCACCATGCTCTGCCACATG 1418
QY 436 ArgProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThr 455
Db 1419 AAG--TTGCGGCTCTTCTACTTCTCGCGGCTGGGTGATCATGACCGTCTTTCATC 1475
QY 456 LeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGlu 475
Db 1476 GCGCTTCTCGCGGAGACCAAGACGTGCCATCGAGGAGATGCTGCTCTGGAAG 1535
QY 476 Gly 476
Db 1536 GGA 1538

RESULT 7
US-09-679-686B-1
; Sequence 1, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: B81160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; PRIORITY FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1675
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1624)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1638)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1668)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1670)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1674)
; OTHER INFORMATION: n = a, c, g or t
US-09-679-686B-1
Alignment Scores: 3.01e-46 Length: 1675
Pred. No.: 532.50
Score: 44.89% Matches: 147
Percent Similarity: Conservative: 90

Best Local Similarity: 27.84% Mismatches: 190
Query Match: 21.67% Indels: 101
Db: 4 Gaps: 13
US-09-886-954A-1 (1-477) x US-09-679-686B-1 (1-1675)
QY 7 GluLutThrGlnProLeuLeuGlyProProGlyGlySerAlaProArgGlyArgArg--- 25
Db 150 GAGGACAGCGAGATGCGCGGGTGTGGCGGGGGGAGCGCGCGCGGGAGCGCGAG 209
QY 26 -----ValPheLeuAlaAlaPheAlaAlaLeuGly 36
Db 210 CTCTACGAGGCGAAGATCATCCGCTACTTCATCTCGCTTGCATCTCGCTCGCGC 269
QY 37 ProLeuSerPheGlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArg 56
Db 270 GGATCCCTCTTCGGCTATGACCTCGAGTCTCAGC----- 305
QY 57 AlaAlaProProAlaProArgLeuAspAlaAlaAlaSerTrpPhe----- 72
Db 306 -----GGCGTGACTTCCATGACGACACTTCTCTGGTGAAGTTCTTCCCGGACGTGTAC 356
QY 72 ----- 72
Db 357 CGGCGAAGCAGCGCACCTGTCAGACGAGACTACTGCAAGTACGACACCAAGTGTCTG 416
QY 73 -----GlyAlaValValThrLeuGlyAlaAlaAla 82
Db 417 ACGTCTTTCACCTCGTCTACTTTCGCGGGCTCGTCTCCACCTTCGCGCGCTCC--- 473
QY 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 102
Db 474 -----TACGTGACCAAGCGCGCGCGCGCGCGCATCATGGGT 515
QY 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
Db 516 GCGCGCGCGAGCTTCTTCGCGCGCGCATCAACGCGCGCGCGCATCAACATCGCCATG 575
QY 123 LeuLeuGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
Db 576 CTATCGTCGACGACATCTCTCTCGCGTGGCTTCGCAATACAGCGCGTGCCT 635
QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 636 GTGTACCTGTGAGATGGCGCGCGCGCTCTCGGGGCGATGCTCAACATCGGTTCAG 695
QY 163 LeuMetValValGlyIleLeuLeuAlaTyrLeu----- 174
Db 696 CTGATGATCACCATCGGCATCTTGGCGCGGAGCTCATCAACTACGGCACCACCAAGATC 755
QY 175 ---AlaGlyTrpValLeuGluTrpArgTrpLeuAlaValLeuGlyCysValProSer 193
Db 756 AAGCGCGGTAC-----GGGTGGCGGTGAGCTGGCGGTGGCGCGGTGGCGCGGCC 809
QY 194 LeuMetLeuLeuMetCysPheMetProGluThrProArgPheLeuLeuThrGlnHis 213
Db 810 ATCATCACCTGGGCTCCCTCTTCTCCGCGACACCCCACTCGCTCTGAGAGGGGCG 869
QY 214 ArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrpGlySerGlu-----Gln 230
Db 870 CACCGGAGGAGGACCGCGCATGCTCGCGCATCGCGCGCACGAGACATCGCGCGAG 929
QY 231 GlyTrpGluAspProIleGlyAlaGluGlnSerPheHisLeuAla----- 246
Db 930 GAGTACGCGGACCTGGTGGCGCGCAGGAGAGGCGCGCGAGTGGCGGCCCGCGGG 989
QY 247 ---LeuLeuArgGlnProGlyIleTyrLys---ProPheIleIleGlyValSerLeuMet 264
Db 990 AACATCTGCGCGCGG-----TACCGCGGCGAGCTCACCATGCGGTGCGATCCCC 1043
QY 265 AlaPheGlnGlnSerGlyValAlaAlaValMetPheTyrAlaGluThrIlePheGlu 284
Db 1044 TTCTTCCAGCAGCTCAAGGGGATCAACGTTCATCATGTTCACGCCCGCGTGTCTTCGAC 1103

```
QY 285 GpualalysPheLysAspSerSer-----LeuAlaSerValValValValValIleGln 302
Db 1104 ACGTGGATTCAAGAAAGAGCCTCTCATGTCTCCGTCTCATCGGGGCGCTCGTCAAC 1163
QY 303 ValLeuPheThrAlaValAlaAlaLeuLeuMetAspArgAlaGlyArgAlaLeuLeu 322
Db 1164 GTCTTCCCAACCGTGTGTCTCATCGTCCAGCGGTCTCGCGCGCGCGCAAGCTGTTTC 1223
QY 323 ValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeu 342
Db 1224 CTCAGGGCGCGCGAGATGATCTGTGCGAGCTCATCTGTGGCGACGCTCATCGCGGCC 1283
QY 343 ThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaGln 362
Db 1284 AAGTTCGGACCGCGCGCGCGG-----1307
QY 363 ProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIle 382
Db 1308 -----GACATCGCAAGGGCTAGCGCGGTCTGTGTCTTCTCATCTCGCGCTAGTC 1361
QY 383 AlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhePro 402
Db 1362 GCGGCTTCCCTCGTGTGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1421
QY 403 LeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPhe 422
Db 1422 CTGGAGATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1481
QY 423 LeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrp 442
Db 1482 TGCATCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1538
QY 443 LeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThr 462
Db 1539 TTCTTCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1598
QY 463 LysGlyLysThrLeuGluGlnIle 470
Db 1599 AAGAACGTGCCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1622

RESULT 8
US-09-679-686B-15
; Sequence 15, Application US/09679686B
; Patent No. 6624343
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Lightner, Jonathan E.
; APPLICANT: Rafalski, J. Antoni
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: HEXOSE CARRIER PROTEINS
; FILE REFERENCE: B01160 US NA
; CURRENT APPLICATION NUMBER: US/09/679,686B
; CURRENT FILING DATE: 2003-01-16
; PRIOR APPLICATION NUMBER: 60/081,131
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: PCT/US99/07561
; PRIOR FILING DATE: 1999-04-07
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 15
; LENGTH: 1960
; TYPE: DNA
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (10)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (28)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
```

```
; NAME/KEY: unsure
; LOCATION: (514)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (533)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (542)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (552)
; OTHER INFORMATION: n = a, c, g or t
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (620)
; OTHER INFORMATION: n = a, c, g or t
US-09-679-686B-15

Alignment Scores:
Pred. No.: 1.33e-44 Length: 1960
Score: 518.00 Matches: 148
Percent Similarity: 47.55% Conservative: 95
Best Local Similarity: 28.96% Mismatches: 198
Query Match: 21.08% Indels: 70
DB: 4 Gaps: 16

US-09-886-954A-1 (1-477) x US-09-679-686B-15 (1-1960)
QY 7 GluGluThrGlnProLeuLeuGlyPro-----ProGly 17
Db 141 GAGAAAAATTCCTCGGTAGGAGGTATTAGCAACGGAGGGGGCAGGAGTACCTTGA 200
QY 18 GlySerAlaProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyPro 37
Db 201 AGCCTCACTCCT-----TTTGTGACGGTAACATGTATAGTTGCAGCATGGGTGG 251
QY 38 LeuSerPheGlyPheAlaLeuGlyTyrSerSerProAlaIle-----ProSerLeu 54
Db 252 TTAATCTTCGGTTACGATATTCGGAATTTCAAGTGGGGTGACATCCATCGATCGTTCTG 311
QY 55 GlnArgAlaAlaProProAlaProArg-----63
Db 312 CTCAGTTTTTCCGTCGGTGTTCGGAAGAAGATTCCGACAAACGGTGAACCACTAC 371
QY 64 -----LeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAla 81
Db 372 TGTCAATACGACAGTCAGACACCTGACGATGTTTACGTCGTGTGTATCTCGCGCGTGTG 431
QY 82 AlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 101
Db 432 CTGTCTGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 491
QY 102 LeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrp 121
Db 492 TTCGGAGGCTTGTCTTCTCCTCNCGGTGCCTTATCAACGGNTTCCCANCAACGTTGG 551
QY 122 MetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAla 141
Db 552 NTGCTCATCGTGGTTCGGATCTTGTCTCGGTTCGGTATCCGGTTCCTCAATCATCTGTG 611
QY 142 ProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysVal 161
Db 612 CCACCTCTANCTATCTGAAATGGCTCCATACAAATATAGAGGAGCATTTGAACATGGCTTT 671
QY 162 GlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu--- 180
Db 672 CAGTTGTCCATCACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 731
QY 181 -----TrpArgTrp-----LeuAlaValLeuGlyCysValProProSerLeuMet 195
```


[illegible]

RESULT 10
 US-09-480-884A-135
 ; Sequence 135, Application US/09480884A
 ; Patent No. 6482597
 ; GENERAL INFORMATION:
 ; APPLICANT: wang, Tongtong
 ; APPLICANT: Pan, Liqun
 ; APPLICANT: Hosken, Nancy A.
 ; APPLICANT: Kalos, Michael D.
 ; APPLICANT: Fanger, Gary R.
 ; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
 ; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
 ; FILE REFERENCE: 210121.455C6
 ; CURRENT APPLICATION NUMBER: US/09/480,884A

Db 936 CGGAGAAAGGTCACCATCTCGAGCTGTCCTCCCGGCTACCGCCAGCCCATC 995
Qy 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
Db 996 CTCATCGTGTGCTGCTGAGCTGTCACGAGCTGCTGCGCATCAACGCTGCTCTAT 1055
Qy 278 TyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerLeuAlaSerValVal 297
Db 1056 TACTCCAGCAGCATCTTCGAGAAGCGGGGTGCAGCAGCCTGTGTATGCCACCATGGC 1115
Qy 298 ValGlyValIleGlnValLeuPheThrAlaValAlaLeuMetAspAlaGly 317
Db 1116 TCCGATATCGTCAACAGCGCTTCACTGTCGTGCTGCTGTTGTGTGAGGAGCAGGC 1175
Qy 318 ArgArgLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
Db 1176 CGGCGAGCCTGCACCTCAT----- 1196
Qy 338 AlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAla 357
Db 1197 -----GGCCTCGCTGCGATGGCGGGTGTGCCATATCATG 1232
Qy 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySer 377
Db 1233 ACATCGCGTAGCATGCTGGAGCAGTACCTGATGTCCTATCTATGAGCATCGTGCC 1292
Qy 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet 397
Db 1293 ATCTTTGGCTTGTGGCTTCTTTGAAGTGGTCTCTGGCCCATCCCATGTTTCATCGTG 1352
Qy 398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn 417
Db 1353 GCTGAATCTTTCAGCAGGCTCCACGTCCAGTGCCTATGCGGTTCAGGCTTCCAC 1412
Qy 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
Db 1413 TGACCTCAAAATTCATTTGGGCGATGCTTCCAGTATGTGGAGCAACTGTGTGTC 1472
Qy 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu----- 453
Db 1473 TAC-----GTCTTCATCATCTTCACTGTGTCTCTGCTGCTGTTCTGTTTC 1511
Qy 454 -----PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThr 471
Db 1512 TTCATCTTCACTTCAAGTTCCTGAGACTAAGCGCGGACCTTCGATGAGATCGCT 1571
Qy 472 AlahisPhe 474
Db 1572 TCGGCTTC 1580

RESULT 11
US-09-542-615A-135
; Sequence 135, Application US/09542615A
; Patent No. 6518256

GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; APPLICANT: Fan, Liqun
; APPLICANT: Kalos, Michael D.
; APPLICANT: Bangur, Chaitanya S.
; APPLICANT: Hosken, Nancy A.
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF LUNG CANCER
; FILE REFERENCE: 210121.455C8
; CURRENT APPLICATION NUMBER: US/09/542.615A
; CURRENT FILING DATE: 2000-04-14
; NUMBER OF SEQ ID NOS: 350
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-542-615A-135

Alignment Scores:
Pred. No.: 5,71e-44 Length: 2856
Score: 514.50 Matches: 145
Percent Similarity: 46.92% Conservative: 91
Best Local Similarity: 28.83% Mismatches: 168
Query Match: 20.94% Indels: 99
DB: 4 Gaps: 14

US-09-886-954A-1 (1-477) x US-09-542-615A-135 (1-2856)

Qy 25 ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeu 44
Db 210 CGCTCATGCTGGCTGTGGGAGGAGCAGTCTTGGTCCCTGCGAGTGTGGCTACAACT 269
Qy 45 Gly----- 45
Db 270 GGATCATCAATGCCCCCAAGAGTGTATCAGAGAGTTCTACAAACAGACATGGGTCCAC 329
Qy 46 ---TyrSerProAlaIleProSerLeuGlnArgAlaAlaProAlaProArgLeu 64
Db 330 CGTATGGGAGAGCATCTGCCACC----- 356
Qy 65 AspAspAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGly 84
Db 357 ---ACGCTCACCGCTCTGCTCCTCTCAGTGGCCATCTTT-----TCTGTGGGGC 407
Qy 85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db 408 ATGATGGCTCTTCTCTGTGGGCTTTTCTGTTAACCGCTTGGCGGGAATTCATG 467
Qy 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
Db 468 CTGATGATGACCTGTGGCTTCTGTCGCGCGTGTCTATGGCTTCTCGAACTG--- 524
Qy 116 AlalaGlnAspValTrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGly 135
Db 525 ---GCGAAGTCTTTGAGATGCTATCTGCGCGCTGTCTATCATCATCGGTGTATCTGGC 581
Qy 136 ValAlaSerLeuValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGly 155
Db 582 CTGACACAGGCTTCGTGCCCATGTATGTGGTGAAGTGTACCCACACAGCCTTTCGTGG 641
Qy 156 LeuLeuGlySerCysValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAla 175
Db 642 GCGCTGGGACCTGTGCACAGCTGGCATCGTCTGCGCATCTCTCATCGCCAGCATGTTTC 701
Qy 176 GlyTrpValLeuGlu-----TrpArgTrpLeuAlaValLeu 187
Db 702 GGC-----CTGGACTCCATCATGGGCAACAGGACCTGTGGCCCTCTGCTGAGCATC 755
Qy 188 GlyCysValProProSerLeuMetLeuLeuMetCysPheMetProGluThrProArg 207
Db 756 ATCTTCATCCCGGCTCTGCTGAGTGCATCGTGTGCTTCTGCGCGAGATCCCGC 815
Qy 208 PheLeuLeu---ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrp 226
Db 816 TTCCTGCTCATCAACCGCAACGAGAGAACCGGGGCTGAGTGTCTAAAGAGTGGC 875
Qy 227 GlySerGluGlnGlyTrpGluAspProIleGlyAlaGluGlnSerPheHis----- 244
Db 876 GGGACAGCTGAGTCAACCATGACCTGCGAGGAGATGAAGGAGAGAGTGGCAGATGATG 935
Qy 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPhe 257
Db 936 CGGAGAGAGAGTCAACCATCTCTGGAGCTGTTCCGCTCCCGCGCTACCGCCAGCCATC 995
Qy 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
Db 996 CTCATCGTGTGCTGCTGAGCTGCTCCAGCAGCTGTCTGGCATCAACGCTCTCTCTAT 1055
Qy 278 TyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerLeuAlaSerValVal 297

Db 1056 TACTCCACGAGCATCTTCGAGAAAGCGGGGTCAGCAGCCTGTGTATGCCACCATGGC 1115
QY 298 ValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGly 317
Db 1116 TCCGGTATCGTCAACACGGCTTCACTGTCGTCTGCTGTTGGTGGAGCAGCAGC 1175
QY 318 ArgArgLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
Db 1176 CCGCGGACCTCGACCTCAT- 1196
QY 338 AlaTyrPheIleValLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAla 357
Db 1197 - 1232
QY 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaIlePheValGlySer 377
Db 1233 ACCATCGCTAGCATGCTGGAGCAGCTACCTGATGCTCTATCTGAGCATCGTGGCC 1292
QY 378 MetCysLeuPheIleAlaGlyPheAlaValAlaGlyTyrGlyProIleProIlePheLeuLeuMet 397
Db 1293 ATCTTGGCTTGTGGCTTCTTGAAGTGGTCTGCGCCCATCCCATGGTTCATCGTG 1352
QY 398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn 417
Db 1353 GGTGAATCTCTACGCGAGGTCACGTCAGTGCATTCGCGGTCGAGGCTTCTCCAAC 1412
QY 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
Db 1413 TGGACCTCAATTCATTTGGGTCATGCTTCCAGTATGTGGAGCACTGTGTGTCCTCC 1472
QY 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu 453
Db 1473 TAC- 1511
QY 454 - 1512
Db 1512 TTCATCTTCACTTCAAAAGTTCCTGAGACTAAAGGCGGACCTTCGATGAGATCGCT 1571
QY 472 AlaHisPhe 474
Db 1572 TCCGGCTTC 1580

RESULT 12

US-09-606-421B-135
; Sequence 135, Application US/09606421B
; Patent No. 6531315

GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong

; APPLICANT: Fan, Liqun

; APPLICANT: Kalos, Michael D.

; APPLICANT: Bangur, Chaitanya S.

; APPLICANT: Hoeken, Nancy

; APPLICANT: Panger, Gary R.

; APPLICANT: Li, Samuel X.

; APPLICANT: Wang, Aijun

; APPLICANT: Skeiky, Yasir A.W.

; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY

; FILE REFERENCE: 210121.455C9

; CURRENT APPLICATION NUMBER: US/09/606,421B

; CURRENT FILING DATE: 2000-06-28

; NUMBER OF SEQ ID NOS: 358

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 135

; LENGTH: 2856

; TYPE: DNA

; ORGANISM: Homo sapien

US-09-606-421B-135

Alignment Scores:

Pred. No.: 5,71e-44 Length: 2856

Score: 514.50 Matches: 145

Percent Similarity: 46.92% Conservative: 91

Best Local Similarity: 28.83% Mismatches: 168
Query Match: 20.94% Indels: 99
DB: 4 Gaps: 14
US-09-886-954A-1 (1-477) x US-09-606-421B-135 (1-2856)
QY 25 ArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeu 44
Db 210 CGCCTCATGCTGGCTGTGGAGGAGCAGTGTGGCTTGGCTCCCTGTCAGTTGGCTACAACT 269
QY 45 Gly- 45
Db 270 GCGATCATCAATGCCCCCGAGAGGTGATCGAGGAGTTCTACACCCAGACATGGGTCCAC 329
QY 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeu 64
Db 330 CCGTATGGGAGAGCATCTGCCACC- 356
QY 65 AspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGly 84
Db 357 ---ACGCTCACACCGCTCTGGTCCCTCTCAGTGGCATCTTT- -TCTGTTGGGGC 407
QY 85 ValLeuGly- -GlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db 408 ATGATGGCTCTCTCTGTGGGCTTTTCGTTAAACCGCTTTGGCGGGGGAATTCATG 467
QY 101 LeuLeuCysSer- -ValProPheValAla- -GlyPheAlaValIleThr 115
Db 468 CTGATGATGAACCTGTGGCTTGTGTCGCCGCTTCATCGGCTTCTCGAAACTG- - 524
QY 116 AlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeuThrGlyLeuAlaCysGly 135
Db 525 ---GGCAAGTCCTTTGAGATGCTGATCTCGGCGCGCTTCATCATCGGTGTGTACTGCGGC 591
QY 136 ValAlaSerLeuValAlaProValTyrIleSerGluIleAlaTyrProAlaValArgGly 155
Db 582 CTGACACACAGGCTTCTGTCGCCCATGTATGTGGGTGAAGTGTACCCACACAGCCTTCTCGTGG 641
QY 156 LeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAla 175
Db 642 GCCCTGGGCGACCTTCACAGCTGGCATCTGTCGGCANTCTCATCCCGCAGGTGTTC 701
QY 176 GlyTrpValLeuGlu- -TrpArgTrpLeuAlaValLeu 187
Db 702 GGC- -CTGGACTCCATCATGCGCAACAGGACCTGTGGCCCTGTCTGTGAGCATC 755
QY 188 GlyCysValProProSerLeuMetLeuLeuMetCysPheMetProGluThrProArg 207
Db 756 ATCTTTCATCCCGGCTGTGTGTCAGTGCATCTGTCGCTTCTGCCCCGAGAGTCCCCGC 815
QY 208 PheLeuLeu- -ThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTrp 226
Db 816 TTCCTGTCTCATCAACCGCAACGAGAGAACCGGGCCAAAGAGTGTCTAAAGAGCTGCGC 875
QY 227 GlySerGluGlnGlyTrpGluAspProProIleGlyAlaGluGlnSerPheHis- - 244
Db 876 GGGACAGCTGAGCTGACCCATCATCTCGAGAGATGAAGAGAGAGTGGCAGATGATG 935
QY 245 - - - - -LeuAlaLeuLeuArgGlnProGlyIleTyrLysProPhe 257
Db 936 CCGGAGAGAAGGTCAACCATCTCTGGAGCTGTTCCGCTCCCCCGCTTACCGCCAGCCCATC 995
QY 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
Db 996 CTCATCGCTGTGCTGAGCTGCCAGCAGCTGTCTGGCATCAACCTGTCTTCTAT 1055
QY 278 TyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValVal 297
Db 1056 TACTCCACGAGCATCTTCGAGAAGCGGGGTCAGCAGCCTGTGTATGCCACCATGGC 1115
QY 298 ValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGly 317
Db 1116 TCCGGTATCGTCAACACAGCCTTCACTGTGTCGTGCTGTTGGTGGAGCAGCAGC 1175

318	Arg	ArgLeuLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly	337
1176	CGCGGACCTGCACCTATA	-----	1196
338	Ala	TyrPheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAla	357
1197	-----GGCCTCGCATGGCGGGTTGTGCCATACTCATG	-----	1232
358	Pro	ValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySer	377
1233	ACCATCGCGTAGCATGCTGGAGCAGTACCCCTGGATGTCCTATCTGAGCATCGTGCC	-----	1292
378	Met	CysLeuPheLeuAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet	397
1293	ATCTTTGGCTTTGTGGCTCTTTTGAAGTGGGTCTCTGGCCCATCCCATGGTTCATCGTG	-----	1352
398	Ser	GluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn	417
1353	GCTGAATCTTCAGCCAGGTCCAGCTCCAGCTGCCATTCCGTTGCAGCGCTTCTCCAC	-----	1412
418	TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro	437	
1413	TGGACCTCAAAATTTTCATTGTGGGCATGTGCTTCAGTATGTGGAGCAACTGTGTGGTCCC	-----	1472
438	Tyr	GlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValIle	453
1473	TAC-----GTCTTCATCATCTTCATGTGCTCCCTGGTTCGTTC	-----	1511
454	-----PheThrLeuPheCysValProGluLuthrLysGlyLysThrLeuGluGlnIleThr	471	
1512	TTTCATCTTCACCTACTTCAAAAGTCTCTGAGACTAAAGCCGGACCTTCGATGAGATCGCT	-----	1571
472	Ala	HisPhe	474
1572	TCGCGCTTC	-----	1580

Qy 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaAalaProAlaProArgLeu 64
Dbb ||||| :|||:
Db 330 CGCTATGGGGAGAGCATCGGCCACC----- 356

Qy 65 AspAspAlaAlaAserTrpPheGlyAlaValThrLeuGlyAlaAlaAalGlyGly 84
Dbb ||||| :|||:
Db 357 ---ACGCTACCAAGCTCTGTCTCCCTCTCAGTGCCATCTTT-----TCGTGTGGGGC 407

Qy 85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Dbb ||||| :|||:
Db 408 ATGATTGGCTCTCTCTGTGGGCCCTTTTCGTTAACCGCTTTGGCCGGCGGAATTAACAATG 467

Qy 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
Dbb ||||| :|||:
Db 468 CTGATGATGAACCTGCTGGCTTCGTCGTGGCGTCTCATGGGCTTCTCGAAACTG--- 524

Qy 116 AlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGly 135
Dbb :::: |::|:
Db 525 ---GGCAAAGTCCTTGAGATGCTGATCCTGGGCGCTTCATCATCGGTGTGTACTCGGC 581

Qy 136 ValAlaSerLeuValAlaProValTyrlIeserGluileAlaTyrrProAlaValArgGly 155
Dbb :::: |::|:
Db 582 CTGACCACAGCCTTCGTCATGATGCGGTAAGTGTACCCACAGCCTTCTCGTGGG 641

Qy 156 LeuLeuGlySerCysValGlnLeuMetValValGlyIleLeuLeuAlaTyrrLeuAla 175
Dbb ||||| :|||:
Db 642 GCCTGGGCACCTGCACACAGCTGGGCATCGTCGCGCATCTCATCGCCACAGGTGTC 701

Qy 176 GlyTrpValLeuGlu-----TrpArgTrpLeuAlaValLeu 187
Dbb ||||| :|||:
Db 702 GGC-----CTGGACTCCATCATGGGCACAACAGGACCTGTGGCCCTCTGTGAGCAT 755

Qy 188 GlyCysValProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrProArg 207
Dbb ::||| :|||:
Db 756 ATCTTCATCCGGCCCTGTCGAGTCATCGTCTGCCCTTCTGCCCGCAGAGTCCCCGC 815

Qy 208 PheLeuLeu---ThrGlnHisArgGgInGluAlaMerAlaAlaAleuArgPheLeuTrp 226
Dbb ||||| :|||:
Db 816 TTCCTGCTCATCAACCCAGAGGAGAACCGGGCCAAGAGTAGTGTCTAAAGAAGCTCGCG 875

Qy 227 GlySerGluGlnGlyTrpGluAspProProIleGlyAlaGluGlnSerPheHis----- 244
Dbb ||||| :|||:
Db 876 GGGACACTGACGTGACCCATGACCTGCAGAGATGAGGAAGAAGAGTGGCAGATGATG 935

Qy 245 -----LeuAlaLeuLeuArgGlnProGlyIleTyrrLysProphe 257
Dbb ||||| :|||:
Db 936 CCGGAGAAGAAGGTCCACATCTGGAGCTGTTCCGCTCCCCTCCGCTACCCCGCAGCCCATC 995

Qy 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValIleAlaValMetPhe 277
Dbb ::||| :|||:
Db 996 CTCATCGCTGTGGTGTGCGAGCTGTCCAGAGCTGTCTGGCATCAACGTGTCTTCTCTAT 1055

Qy 278 TyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValVal 297
Dbb ::||| :|||:
Db 1056 TACTCCAGAGCATCTTCGAGAAGCGGGGTGCAGCAGCTGTGTGTATGCACCATGGC 1115

Qy 298 ValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGly 317
Dbb ::||| :|||:
Db 1116 TCCGGTATCGTCAACACGGCCTTCATGTCGTGCTGTGTGGTGGGGAGCAGCAGGC 1175

Qy 318 ArgArgLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGly 337
Dbb ||||| :|||:
Db 1176 CGGCGGACCTGCACCTCAT----- 1196

Qy 338 AlaTyrrPheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAla 357
Dbb ||||| :|||:
Db 1197 -----GGCTCGCTGGCATGGCGGTGTGCGCATCTCATG 1232

Qy 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySer 377
Dbb ::||| :|||:
Db 1233 ACCATCGCGCTAGACCTGTGGAGCAGCTACCTGGATGTCCTATCTATCTGAGCATCGTGCC 1292

Qy 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProfileProTrpLeuLeuMet 397

```
Db 1293 ATCTTTGGCTTTGTGGCTTCTTTGAAGTGGGTCCTGGGCCCATCCCATGGTTTCATCGTG 1352
QY 398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn 417
Db 1353 GCTGAACCTCTTACGAGGTCACAGTCCAGCTGCAGTTCGCCGTTGCAGGCTTCTCCAAC 1412
QY 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
Db 1413 TGGACCTCAAATTTCAATTGTGGGCATGTGCTTCCAGTATGTGGAGCAACTGTGTGTGCC 1472
QY 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu 453
Db 1473 TAC-----GTCTTCATCATCTTCACTGTGCTCTCTGTTCTGTTTC 1511
QY 454 -----PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluIleThr 471
Db 1512 TTCATCTTCACTTACTTCAAGTTCCTGAGCTAAAGGCCGAGCTTCGATGAGATCGCT 1571
QY 472 AlaHisPhe 474
Db 1572 TCCGGCTTC 1580

RESULT 14
US-09-466-396A-135
; Sequence 135, Application US/09466396A
; Patent No. 6646247
; GENERAL INFORMATION:
; APPLICANT: Wang, Tongtong
; TITLE OF INVENTION: COMPOUNDS AND METHODS FOR THERAPY AND
; FILE REFERENCE: 210121.455C4
; CURRENT APPLICATION NUMBER: US/09/466,396A
; CURRENT FILING DATE: 1999-12-17
; NUMBER OF SEQ ID NOS: 224
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 135
; LENGTH: 2856
; TYPE: DNA
; ORGANISM: Homo sapien
US-09-466-396A-135

Alignment Scores:
Pred. No.: 5,71e-44 Length: 2856
Score: 514.50 Matches: 145
Percent Similarity: 46.92% Conservative: 91
Best Local Similarity: 28.83% Mismatches: 168
Query Match: 20.94% Indels: 99
DB: 4 Gaps: 14

US-09-886-954A-1 (1-477) x US-09-466-396A-135 (1-2856)
QY 25 ArgValPheLeuAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeu 44
Db 210 CGCCTCATGCTGCTGTGGAGGAGCAGTCTTGGCTCCCTGCAGTTTGGCTACACACT 269
QY 45 Gly----- 45
Db 270 GGAGTCATCAATGCCCCAGAGGTGATCGAGGAGTCTTACACAGCATGGGTCCAC 329
QY 46 ---TyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArgLeu 64
Db 330 CGCTATGGGAGAGCATCTGCCACC-----TCTGTGGGGGC 356
QY 65 AspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaGlyGly 84
Db 357 ---ACGCTCACCACGCTGCTGCTCCCTCTCAGTGGCCATCTTT-----TCTGTGGGGGC 407
QY 85 ValLeuGly-----GlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db 408 ATGATGGCTCCTTCTCTGTGGGCCCTTTTCGTTAACCGCTTTGGCCGGCGGAATTCATG 467
QY 101 LeuLeuCysSer---ValProPheValAla-----GlyPheAlaValIleThr 115
```

```
Db 468 CTGATGATGAACCTGTGGCTTCTGTGTCGGCGTCTCATGGGCTTCTCGAAATCG--- 524
QY 116 AlaAlaGlnAspValTrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGly 135
Db 525 ---GGCAAGTCTTTGAGATGCTGATCTCTGGGCGGCTTCTCATCTCGGTGTACTCGGC 581
QY 136 ValAlaSerLeuValAlaProValTyrIleSerGluIleAlaLafyrProAlaValArgGly 155
Db 582 CTGACCACAGGCTTCGTGCCCATGTATGTTGGGTGAAGTGTCAACACAGCTTTCTGGG 641
QY 156 LeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAlaLafyrLeuAla 175
Db 642 GGCCTGGGACCTGTGCACAGCTGGGCATCTGTGCGCATCTCTATCGCCAGGTGTTTC 701
QY 176 GlyTrpValLeuGlu-----TyrArgTrpLeuAlaValLeu 187
Db 702 GGC-----CTGACTCCATCATGGGCAACAGGAGCTGTGGCCCTCTGTCTGAGCATC 755
QY 188 GlyCysValProSerLeuMetLeuLeuMetCysPheMetProGluThrProArg 207
Db 756 ATCTTCATCCGGCCCTGCTGCAGTGCATCTGTGCTCCCTTCTGCCCCGAGAGTCCCCG 815
QY 208 PheLeuLeu---ThrGlnHisArgGlnGlnAlaMetAlaAlaLeuArgPheLeuTrp 226
Db 816 TTCCTGCTCATCAACCGCAACGAGGAGAACCGGGCCAGAGTGTCTAAAGAAGCTCGCG 875
QY 227 GlySerGluGlnGlyTrpGluAspProPheGlyAlaGluGlnSerPheHis----- 244
Db 876 GGGACAGCTGACGTGACCATCCCTGCAGAGATGAAGGAGAGAGTCCGAGATGATG 935
QY 245 -----LeuAlaLeuLeuAsgGlnProGlyIleTyrLysProPhe 257
Db 936 CGGGAGAGAGAGTCAACCATCTGTGAGTGTCCGCTCCCGCTCCCGCCACGAGCCATC 995
QY 258 IleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPhe 277
Db 996 CTCATCGCTGTGCTGCTGAGCTGTCCACAGAGTGTCTGGCATCAACGCTGTTCTAT 1055
QY 278 TyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValVal 297
Db 1056 TACTCCAGAGCATCTTCGAGAAGCGGGGGTGCAGACGCTGTGTATCCACCATTTGGC 1115
QY 298 ValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGly 317
Db 1116 TCGGTATCGTCAACACGGCTTCACTGTCTGTGCTGCTGTTGTGTGGTGGAGCGAGGCG 1175
QY 318 ArgArgLeuLeuLeuValSerGlyValValMetValPheSerThrSerAlaPheGly 337
Db 1176 CGGCGGACCTGCACCTCAT----- 1196
QY 338 AlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAla 357
Db 1197 -----GGCCTCGCTGGCATCGCGGTTGTGCCATCTACTCATG 1232
QY 358 ProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySer 377
Db 1233 ACATCGCGCTAGCAGTGTGGAGCAGTACCTCGGATGCTCTATCTGAGCATCTGGGCC 1292
QY 378 MetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMet 397
Db 1293 ATCTTTGGCTTTGTGGCTTCTTTGAAGTGGGTCTCTGGCCCATCCCATCGTTCATCGTG 1352
QY 398 SerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsn 417
Db 1353 GCTGAACCTTCAAGCTTTCAGCGAGGTCACAGTCCAGCTGCCATTTGGCGGCTTCTCCAAC 1412
QY 418 TrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgPro 437
Db 1413 TGGACCTCAAATTTCAATTGTGGGCATGTGCTTCCAGTATGTGGAGCAACTGTGTGTGCC 1472
QY 438 TyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu----- 453
```


Search completed: February 25, 2005, 21:35:54
Job time : 276 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 25, 2005, 20:31:27 ; Search time 677 Seconds
(without alignments)
4174.955 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457

Sequence: 1 MTPEDPEETQPLPGPGSSA.....CVPETKGTLEQITAFHEGR 477

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 5394803 seqs, 2962729879 residues

Total number of hits satisfying chosen parameters: 10789606

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DSV=xlp
-Q=/cgn2_1/USPTO.spool_p/US09866954/runat_23022005_154244_13316/app_query.fasta_1.647
-DB=Published Applications NA -QFMT=fastap -SURF=p2n.rnpb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -LOCAL=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=200000000 -USER=US09866954 @CGN 1 1 723 @runat_23022005_154244_13316
-NCFU=6 -ICPU=3 -NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSFBLOCK=100
-LONGLOG -DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-XGAPOP=6 -XGAPEXT=7 -XGAPOP=10 -XGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

- 1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq*
- 2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq*
- 3: /cgn2_6/ptodata/1/pubpna/US06_NEW_PUB.seq*
- 4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq*
- 5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq*
- 6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq*
- 7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq*
- 8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq*
- 9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq*
- 10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq*
- 11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq*
- 12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq*
- 13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq*
- 14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq*
- 15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq*
- 16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq*
- 17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq*
- 18: /cgn2_6/ptodata/1/pubpna/US10F_PUBCOMB.seq*
- 19: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq*
- 20: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq*
- 21: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq*
- 22: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq*

pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2457	100.0	1862	17	US-10-115-831-137 Sequence 137, Appl
2	2457	100.0	2080	16	US-10-168-651-28 Sequence 28, Appl
3	2067.5	84.1	1655	17	US-10-115-831-138 Sequence 138, App
4	1679.5	68.4	1156	17	US-10-264-237-688 Sequence 688, App
5	1554.5	63.3	1461	17	US-10-169-395-29 Sequence 19, Appl
6	1268	51.6	789	17	US-10-169-395-19 Sequence 29, Appl
7	958.5	39.0	2356	15	US-10-157-031-358 Sequence 358, App
8	958.5	39.0	2487	17	US-10-172-118-1685 Sequence 1685, App
9	958.5	39.0	2487	17	US-10-342-887-1685 Sequence 1685, App
10	958.5	39.0	2487	18	US-10-755-889-109 Sequence 109, App
11	958.5	39.0	2696	10	US-09-814-353-19608 Sequence 19608, A
12	716	29.1	2202	17	US-10-424-599-132422 Sequence 132422, A
13	702	28.6	1797	17	US-10-425-114-26596 Sequence 26596, A
14	702	28.6	2003	18	US-10-437-963-61061 Sequence 61061, A
15	701.5	28.6	2049	17	US-10-425-114-24968 Sequence 24968, A
16	701.5	28.6	2244	18	US-10-425-115-138759 Sequence 138759, A
17	684	27.8	1828	18	US-10-739-930-1250 Sequence 1250, Ap
18	682	27.8	1875	17	US-10-425-114-26550 Sequence 26550, A
19	666.5	27.1	1953	13	US-10-051-909-35 Sequence 35, Appl
20	662.5	27.0	1963	18	US-10-739-930-1789 Sequence 1789, Ap
21	639	26.0	1768	17	US-10-310-154-357 Sequence 357, App
22	620.5	25.3	150130	13	US-10-087-192-820 Sequence 820, App
23	609.5	24.8	1785	18	US-10-437-963-64291 Sequence 64291, A
24	607	24.7	1386	17	US-10-369-493-47058 Sequence 47058, A
25	604.5	24.6	2045	17	US-10-425-114-22342 Sequence 22342, A
26	604.5	24.6	2051	17	US-10-425-114-29817 Sequence 29817, A
27	604.5	24.6	2190	18	US-10-437-963-74980 Sequence 74980, A
28	603	24.5	1419	17	US-10-369-493-47223 Sequence 47223, A
29	598.5	24.4	1395	17	US-10-282-122A-20182 Sequence 20182, A
30	593.5	24.2	1766	17	US-10-425-114-22366 Sequence 22366, A
31	593.5	24.2	1767	18	US-10-437-963-22036 Sequence 22036, A
32	591.5	24.1	1395	17	US-10-369-493-47239 Sequence 47239, A
33	591.5	24.1	1395	17	US-10-282-122A-6734 Sequence 6734, Ap
34	591.5	24.1	1446	9	US-10-812-315A-3 Sequence 3, Appl
35	587.5	23.9	1359	9	US-09-938-842A-1627 Sequence 1627, Ap
36	587.5	23.9	1359	11	US-09-938-842A-1627 Sequence 1627, Ap
37	584.5	23.8	1473	9	US-09-938-842A-2331 Sequence 2331, Ap
38	584.5	23.8	1473	11	US-09-938-842A-2331 Sequence 2331, Ap
39	583.5	23.7	1395	17	US-10-282-122A-38732 Sequence 38732, A
40	583.5	23.7	1395	17	US-10-282-122A-40124 Sequence 40124, A
41	582	23.7	1566	9	US-09-938-842A-1769 Sequence 1769, Ap
42	582	23.7	1566	11	US-09-938-842A-1769 Sequence 1769, Ap
43	582	23.7	1898	17	US-10-424-599-96711 Sequence 96711, A
44	574.5	23.4	1673	18	US-10-664-705-79 Sequence 79, Appl
45	571	23.2	1338	17	US-10-282-122A-35021 Sequence 35021, A

ALIGNMENTS

RESULT 1
US-10-115-831-137
; Sequence 137, Application US/10115831
; Publication No. US20030219743A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyao
; APPLICANT: Ren, Feiyao
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219743A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 792CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/115,831
; CURRENT FILING DATE: 2002-04-02
; PRIOR APPLICATION NUMBER: 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 178


```

; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 137
; LENGTH: 1862
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)..(1437)
US-10-115-831-137

```

```

Alignment Scores:
Pred. No.: 1,78e-241 Length: 1862
Score: 2457.00 Matches: 477
Percent Similarity: 100.00% Conservativeness: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 17 Gaps: 0

```

US-09-886-954A-1 (1-477) x US-10-115-831-137 (1-1862)

```

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
DB 4 ATGACGCCGAGGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGGGCGCAGCGC 63
QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
DB 64 CCCCGCGCGCGCGCTTCTCGCGCGCTTCTCGCGCGCTTCTCGCGCGCTTCTCGCGCT 123
QY 41 GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
DB 124 GGTTCGCGCTCGCTACAGCTCCCGCGCATCCCTAGCGCTGACGCGCGCGCGCGCG 183
QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
DB 184 GCCCGCGCGCTGACGACGCGCGCGCTCTGTTCTGGGGCTGTCTGACCTCGGTCGC 243
QY 81 AlaAlaGlyClyValLeuGlyTrpLeuValAspArgAlaGlyValGlyLeuSerLeu 100
DB 244 GCGCGCGGGGAGTGTCTGGCGCGCTGCTGTGTGAGCCGCGCGCGCGCGCGCGCGCT 303
QY 101 LeuLeuGlySerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
DB 304 TTGCTGTGCTCGTGTGCTGCGCGCGCTTCTGCGCGCTTCTGCGCGCTTCTGCGCG 363
QY 121 TrpMetLeuGlyGlyArgLeuLeuThrGlyValAlaCysGlyValAlaSerLeuVal 140
DB 364 TGGATGTCTGGGGGCGCGCTTCTGACCGCGCTTCTGCGCGCTTCTGCGCGCTTCTG 423
QY 141 AlaProValTyrLeuSerGluAlaAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
DB 424 GCCCGCGCTTACATCTCCGAATCGCTACCGAGCAGTCCGGGGTGTCTGCGCTCTCTGT 483
QY 161 ValGlnLeuMetValValValGlyLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
DB 484 GTGCAGCTAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 543
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 544 TGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 603
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
DB 604 TTCATGCCGAGACCCCGCTTCTGCTGACTCAGCACAGGCGCCAGGAGGCGCATGGCC 663
QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheGlyAlaGlu 240
DB 664 GCCCTGCGGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 723
QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLeuProPheIleLeuGly 260
DB 724 CAGAGCTTTCACCTGGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 783
QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280

```

```

DB 784 GTCTCCTCATGGCCTTCCAGCAGCTGTCTGGGGGTCAAGCCGCTCATGTTCTATGCAGAG 843
QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyVal 300
DB 844 ACCATCTTTGAAGAGGCCAAGTTCAAGGACAGCAGCTGGCTCGGTCTGTGGGTGTCTC 903
QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
DB 904 ATCCAGGTGTCTTTCAGCTGTGGCGGCTCTCATCATGACAGAGCAGGCGGAGGCTGTG 963
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
DB 964 CTCTGTGTCTTGTTCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1023
QY 341 LysLeuThrGlnGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSer 360
DB 1024 AAGCTGACCCAGGCTGGCCCTGCGCACTCTCTGACAGTGGCCATCTCGGCGCTGTCTCT 1083
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
DB 1084 GCACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1143
QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProIleProIleLeuMetSerGluIle 400
DB 1144 TTCTATCGCGCTTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1203
QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMet 420
DB 1204 TTCCCTCTGATGTCAAGGCGGTGGGACAGGCACTGTGTGTCTTCCACCACTGCTCATG 1263
QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
DB 1264 GCCTTTCTGTGACCAAGGAGTTTCAGCAGCTCATGAGGCTCTCAGGCGCTATGGAGCC 1323
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
DB 1324 TTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1383
QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
DB 1384 GAACTAAGAAAGACTCTGGAACAAATCACAGCCCAATTTGAGGGGCGCA 1434

```

RESULT 2

```

US-10-168-651-28
; Sequence 28, Application US/10168651
; Publication No. US20030171275A1
; GENERAL INFORMATION:
; APPLICANT: INCTE GENOMICS, INC.
; APPLICANT: BAUGHN, Marian R.
; APPLICANT: BURFORD, Neil
; APPLICANT: AU-YOUNG, Janice
; APPLICANT: LU, Dying Aina M.
; APPLICANT: YANG, Junming
; APPLICANT: REDDY, Roopa
; APPLICANT: LAL, Preeti
; APPLICANT: HILLMAN, Jennifer L.
; APPLICANT: AZIMZAI, Yalda
; APPLICANT: YUE, Henry
; APPLICANT: NGUYEN, Daniel B.
; APPLICANT: YAO, Monique G.
; APPLICANT: GANDHI, Ameeni R.
; APPLICANT: TANG, Y. Tom
; APPLICANT: KHAN, Farrah A.
; TITLE OF INVENTION: TRANSPORTERS AND ION CHANNELS
; FILE REFERENCE: PI-0005 PCT
; CURRENT FILING DATE: 2002-06-21
; PRIOR APPLICATION NUMBER: US/10/168,651
; PRIOR FILING DATE: 2002-06-21
; PRIOR FILING DATE: 1999-12-23; 2000-01-14; 2000-01-21; 2000-01-28; 2000-02-02;
; 2000-02-10
; NUMBER OF SEQ ID NOS: 54

```

```

; SOFTWARE: PERL Program
; SEQ ID NO 28
; LENGTH: 2080
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20030171275A1 1416107CB1
US-10-168-651-28

Alignment Scores:
Pred. No.: 2,06e-241 Length: 2080
Score: 2457.00 Matches: 477
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 16 Gaps: 0

US-09-886-954A-1 (1-477) x US-10-168-651-28 (1-2080)

QY 1 MetThrProGluAspProGluGlnThrGlnProLeuLeuGlyProGlyGlySerAla 20
DB 46 ATGACGCCCGAGGACCCAGAGGAACCCAGCCGCTTCTGGGGCTCTCTGGGGCGCGG 105
QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
DB 106 CCGCGGGCGCGCGCTTCTCTCGCGCGCTTCTCGCGCGCTTCTCGCGCGCTTCTCG 165
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaProPro 60
DB 166 GGGTTCCGCTCGCTACAGCTCCCGCGCATCCCTAGCTGCGAGCGCGCGCGCGCG 225
QY 61 AlaProArgLeuAspAspAlaAlaSerTrpPheGlyValValThrLeuGlyAla 80
DB 226 GCCCGCGCTGACACGCGCGCGCTCTCTGGTTCGGGGCTGCTGACCTTGGGTGCC 285
QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSer 100
DB 286 GCGGGCGGGGAGTGTGGCGGCTGGCTGGTGAGCGCGCGCGCGCAAGCTGAGCCTC 345
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaGlnAspVal 120
DB 346 TTGCTGTCTCGCTCGCTTCTGGTGGCGGCTTTCGCGCTCATCACCGCGCGCGGAG 405
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
DB 406 TGGATGCTGTGGGGGGCGGCTCTCACGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 465
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
DB 466 GCCCGGCTTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGTGTCTCGGCTCTGT 525
QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrIleuAlaGlyTrpValLeuGlu 180
DB 526 GTGACGTAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 585
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 586 TGGCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 645
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
DB 646 TTCATGCCCGAGACCCCGCGCTTCTGTGCTGACTCAGCAGAGGGGTGGAGGCCATGG 705
QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheIleGlyAlaGlu 240
DB 706 GCCCTCGGTTCTGTGTGGGCTCCGAGCAGGGGTGGAGAGACCCCGCCCATCGGGCTG 765
QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleYrsProPheIleGly 260
DB 766 CAGAGCTTTACCTGGCGCTGTGTGGCGAGCGCGGCTCTTACAAAGCCCTTCATCATCG 825
QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
;
826 GTCTCCCTCATGGCTTCCAGCAGCTGTGGGGGTCAACGCCGTCTATGTCTATGCAGAG 885
QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyVal 300
DB 886 ACCATCTTTGAAGAGGCCAAGTTCAAGGACAGCAGCTGGCTCGGTCTGGTGGGTGTC 945
QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgLeu 320
DB 946 ATCCAGGTGTGTTCACAGCTGTGGCGGCTCTCATCATGACAGACAGCGGGGAGGCTG 1005
QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
DB 1006 CTCCTGGTCTTTCAGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1065
QY 341 LysLeuThrGlnGlyGlyProGlyLysSerHisValAlaIleSerAlaProValSer 360
DB 1066 AAGCTGACCCAGGGTGGCGCTTGGCACTCTCGACGTGGCCATCTCGGGCGCTGTCT 1125
QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
DB 1126 GCACAGCCTGTGTGATGCCAGCGTGGGGCTGGCTGGCTGGCTGGCGAGCATGTGCTC 1185
QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400
DB 1186 TTCTGCGCGGCTTTCGCGTGGGGTGGGGCGCCATCCCTGGCTCTCATGTGCAGAGATC 1245
QY 401 PheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrLeuMet 420
DB 1246 TTCCCTCTCATGTCAAGGGCGTGGCGAGCGGATCTGGCTCTCTCAACACTGGCTCATG 1305
QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
DB 1306 GCCTTCTGTGACCAAGGAGTTTCAGCAGCCTCATGGAGGTCTCATGGCCCTATGGAGCC 1365
QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
DB 1366 TTCTGGCTGGCTCGGCTTCTGCTATCTTCAGTGTCTTCTTCTTCTTCTTCTTCTTCT 1425
QY 461 GluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
DB 1426 GAACTAAAGGAAGAACTCTGGAACAAATCACAGGCCATTTTGAGGGGGGA 1476

RESULT 3
US-10-115-831-138
; Sequence 138, Application US/10115831
; Publication No. US20030219743A1
; GENERAL INFORMATION:
; APPLICANT: Tang, Y. Tom
; APPLICANT: Liu, Chenghua
; APPLICANT: Asundi, Vinod
; APPLICANT: Ren, Feiyan
; APPLICANT: Drmanac, Radoje T.
; TITLE OF INVENTION: No. US20030219743A1el Nucleic Acids and
; TITLE OF INVENTION: Polypeptides
; FILE REFERENCE: 792CIP2ADIV
; CURRENT APPLICATION NUMBER: US/10/115,831
; PRIOR APPLICATION NUMBER: 09/667,298
; PRIOR FILING DATE: 2000-09-22
; PRIOR APPLICATION NUMBER: 09/577,408
; PRIOR FILING DATE: 2000-05-18
; NUMBER OF SEQ ID NOS: 178
; SOFTWARE: pt_FL_genes Version 2.0
; SEQ ID NO 138
; LENGTH: 1655
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (4)..(1230)
US-10-115-831-138
```

Alignment Scores:		1.26e-201	Length: 1655
Pred. No.:	2067.50	Matches: 408	
Percent Similarity:	85.53%	Conservative: 0	
Best Local Similarity:	85.53%	Mismatches: 0	
Query Match:	84.15%	Indels: 69	
DB:	17	Gaps: 1	
US-09-886-954A-1 (1-477) x US-10-115-831-138 (1-1655)			
QY	1	MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla	20
DB	4	ATGACCCCGAGGACCCAGAGAAACCCAGCCGCTTCGTGGGCCCTTCCTGGCGGAGCGCG	63
QY	21	ProArgGlyArgValPheLeuAlaAlaPheAlaAlaLeuLeuGlyProLeuSerPhe	40
DB	64	CCCCGGCGCGCGCTTCCTCGCGGCTTCGCGGCTTCGCGGCTTCGCGGCTTCGCGGCTTC	123
QY	41	GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro	60
DB	124	GGCTTCGCGCTCGGCTACAGCTCCCGGCGCATCCCTAGCTGACAGCGCGCGCGCCCG	183
QY	61	AlaProArgLeuAspAlaAlaAlaSerTTPheGlyAlaValValThrLeuGlyAla	80
DB	184	GGCCCGCGCTGGACGCGCGCGCTTCCTGCTGCTGCGG	222
QY	81	AlaAlaGlyGlyValLeuGlyGlyTTPLeuValAspArgAlaGlyArgLysLeuSerLeu	100
DB	222	-----	222
QY	101	LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal	120
DB	222	-----	222
QY	121	TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal	140
DB	222	-----	222
QY	141	AlaProValTyrLeuSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys	160
DB	223	-----GTCTACATCTCCGAAATCGCTACCCAGCAGTCCGGGGTTCCTCGGCTCCTGT	276
QY	161	ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTTPValLeuGlu	180
DB	277	GTGCAGCTAATGTGTCGTCGCGCATCTCTGCGCTACCTGCAGCGCTGGTGTGGAG	336
QY	181	TrpArgTTPLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys	200
DB	337	TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG	396
QY	201	PheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGlnAlaMetAla	220
DB	397	TTCATGCCCGAGACCCCGCTTCCTGCTGACTCAGCACAGGCGCCAGAGGCCATGGCC	456
QY	221	AlaLeuArgPheLeuTTPGlySerGluGlnGlyTTPGluAspProPheIleGlyAlaGlu	240
DB	457	GCCCTGCGCTTCCTGTGGGCTCCGAGCAGGCTGGGAGACCCGCCATCGGGGTGAG	516
QY	241	GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleGly	260
DB	517	CAGAGCTTTCACCTGGCGCTGCTGGCGAGCCGCGCATCTACAGCCCTTCATCATCGGT	576
QY	261	ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu	280
DB	577	GTCTCCCTGATGGCTTCAGAGAGCTGTCGGGGTTCACGCGCTCATGTTCATGACAG	636
QY	281	ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal	300
DB	637	ACCATCTTGAAGAGCCCAAGTTCAAGGACAGCAGCCCTGGCTCGGTGCTGGTGTGTC	696
QY	301	IleGlnValLeuPheThrAlaValAlaIleLeuMetAspArgAlaGlyArgLeu	320
DB	697	ATCCAGGTGCTGTTACAGCTGTGGCGGCTCTCATCATGGACAGAGCGGCGAGGCTG	756

RESULT 4
US-10-264-237-688
; Sequence 688, Application US/10264237
; Publication No. US20040009491A1
; GENERAL INFORMATION:
; APPLICANT: Birst et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: P131P1
; CURRENT APPLICATION NUMBER: US/10/264,237
; CURRENT FILING DATE: 2002-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/16450
; PRIOR FILING DATE: 2001-05-18
; PRIOR APPLICATION NUMBER: US 60/205,515
; PRIOR FILING DATE: 2000-05-19
; NUMBER OF SEQ ID NOS: 2876
; SOFTWARE: Patent In Ver. 3.1
; SEQ ID NO 688
; LENGTH: 1156
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (24)..(24)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (47)..(47)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (351)..(351)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (397)..(397)
; OTHER INFORMATION: n equals a,t,g, or c
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (432)..(432)
; OTHER INFORMATION: n equals a,t,g, or c


```

; LENGTH: 1461
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (61)..(849)
US-10-169-395-29

Alignment Scores:
Pred. No.: 46-149 Length: 1461
Score: 1554.50 Matches: 324
Percent Similarity: 67.86% Conservative: 1
Best Local Similarity: 67.64% Mismatches: 1
Query Match: 63.27% Indels: 153
DB: 17 Gaps: 2

US-09-886-954A-1 (1-477) x US-10-169-395-29 (1-1461)
QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
DB 61 ATGACCGCCGAGACCCAGAGAAACCCAGCGCTTCTGGGGCCCTCTGGCGGCGAGCGG 120
QY 21 ProArgGlyArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPhe 40
DB 121 CCCCAGCGCGCGCGCTTCTCTCGCGCGCTTCTCGCGCTGCGCTGGGCCACTCAGCTTC 180
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
DB 181 GGCTTCGCGCTCGGCTACAGCTCCCGCGCATCCCTAGCTGCAGCGCGCGCGCCCGG 240
QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
DB 241 GCCCGCGCGCTGAGACACCGCGCGCTCTCTGTTCCGGGCTGTGTCGACCCCTGGTGCC 300
QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgIysLeuSerLeu 100
DB 301 CGCGCGGGGGAGTGTGGCGGCTGGTGGTGGACCGCGCGCGCGCAAGCTGAGCCTC 360
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaGlnAspVal 120
DB 361 TTGCTGTGCTCCGCTTCGTGGCGGCTTTCGCTCATCCCGCGCGCGCGAGACGTG 420
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
DB 421 TGGATGCTGTGGGGGCGCGCTCTCCACCGGCTGGCTGGGTGCTCCCTCCCTAGTG 480
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
DB 481 GCCCGGCTCTACATCTCCGAATCGCTACCCAGCAGTCCGGGGTGTGCTCGGCTCCTGT 540
QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrIleuAlaGlyTrpValLeuGlu 180
DB 541 GTGCAGCTAATGTGTGTCGCGCATCTCTCGGCTACCTGGCAGGCTGGGTGGTGGAG 600
QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
DB 601 TGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 660
QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAla-MetAl 220
DB 661 TTCATGCCCGAGACCGCGCTTCTGCTGACTCAGCAGCAGGCGCGAGGGGTGCTC-- 718
QY 220 aAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaG 240
DB 718 ----- 718
QY 240 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrIleProPheIleleG 260
DB 718 ----- 718
QY 260 yValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaG 280
DB 718 ----- 718
```

```

QY 280 uThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVa 300
DB 718 ----- 718
QY 300 lIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLe 320
DB 718 ----- 718
QY 320 uLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPh 340
DB 719 -----CTGCTCTTGTGAGGTGTGTCATGGTGTTCAGCACGAGTGCCTTCGGCGCCACTT 774
QY 340 eLysLeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSe 360
DB 775 CAAGCTGACCCAGGTTGGCCCTTGGCACTCTCTGCACGTGGCCATCTCGGCGGCTGTCTC 834
QY 360 rAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLe 380
DB 835 TGCACAGCCTGTTGATGCCAGCGTGGGCTGGGCTGGCTGGCGTGGGCGACATGTGCTC 894
QY 380 uPheIleAlaGly-PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluI 400
DB 895 CTTTCATCGCGGAGGTC----- 911
QY 400 lePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuM 420
DB 911 ----- 911
QY 420 etAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyA 440
DB 912 -----CTCAGGCGCTATGGAG 927
QY 440 laPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValP 460
DB 928 CCTTCTGCTTGGCTTCCGCTTCTGCACTTTCAGTGTCTTTCCTTTCACCTTGTCTGTCTCC 987
QY 460 roGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
DB 988 CTGAACATAAGGAAGAACTCTGGAACAATCAGACCCATTTTGAGGGGCGA 1040

RESULT 6
US-10-169-395-19
; Sequence 19, Application US/10169395
; Publication No. US20040034192A1
; GENERAL INFORMATION:
; APPLICANT: KATO, Seishi
; APPLICANT: KIMURA, Tomoko
; TITLE OF INVENTION: HUMAN PROTEINS HAVING HYDROPHOBIC DOMAINS AND DNAS ENCODING
; TITLE OF INVENTION: THESE PROTEINS
; FILE REFERENCE: 01997.015100. US
; CURRENT APPLICATION NUMBER: US/10/169,395
; CURRENT FILING DATE: 2002-11-29
; PRIOR APPLICATION NUMBER: JP 2000-585
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-588
; PRIOR FILING DATE: 2000-01-06
; PRIOR APPLICATION NUMBER: JP 2000-2299
; PRIOR FILING DATE: 2000-01-11
; PRIOR APPLICATION NUMBER: JP 2000-26862
; PRIOR FILING DATE: 2000-02-03
; PRIOR APPLICATION NUMBER: JP 2000-58367
; PRIOR FILING DATE: 2000-03-03
; PRIOR APPLICATION NUMBER: PCT/JP00/09359
; PRIOR FILING DATE: 2000-12-28
; NUMBER OF SEQ ID NOS: 150
; SEQ ID NO 19
; LENGTH: 789
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-169-395-19
Alignment Scores:
```


Db 1493 ACGGTCCTCGGAGCAGATCGAGTCCTCTTCTCGCAGCGGGAGA 1535

```

432 CTGCTGCTCGGAGGACGCTGAGCGGCTTCGCCGGGGGCTCACAGCTGCTGCATCCG 491
143 ValTyrIleSerGluIleAalaTyrPrAlaValArgGlyLeuLeuGlySerCysValGln 162
492 GTGTACGTCTGTAGATTGCTCCCGCAGCGCTGCTGGGGCTCTGGGGGGCCACACCCAG 551
163 LeuMetValValGlyIleLeuLeuAalaTyrIleLeuAalaGlyTyrValLeuGluTyrArg 182
552 CTCATGGCAGTGTGGATCCCTGCTCTTACGCCCTTGGCTCTGCTGCTGCTGCGGCG 611
183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
612 TGGCTGGCTGTGGCGGGGAGCGCTGTGCTCATCATGATCTGCTGCTGCTGCTGCTCATG 671
203 ProGluThrProArgPheLeuThrGlnHisArgArgGlnGlnAlaMetAlaAlaLeu 222
672 CCCAACTCGCGGCTTCTGCTCTCTCGGGCAGGAGAGAGAGAGAGAGAGAGAGAGAG 731
223 ArgPheLeuTyrGlySerGluGlnGly-----TrpGluAspProProIleGly----- 238
732 GCTGTGCTGGTGGGAGCGAGCTGATGTCCTGAGTTCGAGCAGATCCAGGACAAAC 791
239 ---AlaGluGlnSerPheHisLeu-----AlaLeuLeuArgGlnProGlyIleTyrIys 255
792 GTCGGAGACAGACGACGAGTATCTGGGCTGAGGACGAGCGGCGGCGGCGGCGGCGG 851
256 ProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaVal 275
852 CCCATCACCTGCGCTGTGATGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 911
276 MetPheTyrAlaGluThrIlePheGluGluAlaLysPhe-----LysAspSer 291
912 CTGCTGTACCTGCGAGCGCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 971
292 SerLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeu 311
972 -----GCAGCATCTGTGGGGCGTGGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1022
312 IleMetAspArgAlaGlyArgArgLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 331
1023 ACCATGGACCTCGCAGCGCGCAAGTGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1082
332 SerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSer 351
1083 GCCAACCTGACTCTGGGGCTGTATCCACTTT-----GCCCGCAGGCTCTGAGC 1133
352 HisValAlaIleSerAlaProValSer-----AlaGlnProValAsp 365
1134 CCCAACGACCTGCGGCGCTGGAAGCGAGTCTGCTGGGGGAGCTTGGCGCAGCGCTGGCA 1193
366 AlaSerValGly---LeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGly 384
1194 GCACCGCTGCTACCTACCTCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1253
385 PheAlaValGlyTyrGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHis 404
1254 TACGCGGTGGGTGGGGTCCCATCACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1313
405 ValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424
1314 GCCCGTGGGTGGCTCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTC 1373
425 ThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
1374 ACCAAGTCTCTCTCCGACGAGTGT-GAGCACCTTCGGCTCCAGGCTGCTGCTGCTGCTG 1432
445 -SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLysG 464
1433 CGCGGCGCATCTGCTGGTGGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1492
464 yLysThrLeuGluGlnIleThrAlaHisPheGlu---GlyArg 477

```

```

RESULT 8
US-10-172-118-1685
; Sequence 1685, Application US/10172118
; Publication No. US20030224374A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Chris
; APPLICANT: Van 't Veer, Laura
; APPLICANT: Van de Vijver, Marc
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-175-999
; CURRENT APPLICATION NUMBER: US/10/172,118
; CURRENT FILING DATE: 2002-06-14
; PRIOR APPLICATION NUMBER: 60/380,770
; PRIOR FILING DATE: 2002-05-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1685
; LENGTH: 2487
; TYPE: DNA
; ORGANISM: Homo sapiens
; PUBLICATION INFORMATION:
; DATABASE ACCESSION NUMBER: NM_017585
; DATABASE ENTRY DATE: 2001-06-18
US-10-172-118-1685

Alignment Scores:
Pred. No.: 9,86e-88 Length: 2487
Score: 958.50 Matches: 227
Percent Similarity: 59.22% Conservative: 78
Best Local Similarity: 44.08% Mismatches: 164
Query Match: 39.01% Indels: 47
DB: 17 Gaps: 12

US-09-886-954A-1 (1-477) x US-10-172-118-1685 (1-2487)
QY 3 ProGlu---AspProGluGluThrGlnProLeuLeuGly----- 14
Db 20 CCCGAGAGAGACCCGGCCATGCGAGAGCCGCTGCTGGAGCCGCGGAGCTACGAC 79
QY 15 -----ProProGlyGlySerAlaProArgGlyArg----- 24
Db 80 ACCTTCCCGAGAGAGCCGCCCTGCGCAGGGGACAGGGCGGGTCCGGACCCCTGCAG 139
QY 25 -----ArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPheGlyPhe 42
Db 140 AACAAAGGGTGTCTTCTGGCCACCTTCCGCGCAGTGTCCGGCAATTTTCAGCTTTGGGTAT 199
QY 43 AlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaPro 62
Db 200 GCGCTGGTGTACACATCCCTGTATCCAGCCCTGAGGCGCTCTCTGGATCTCTGAGCTG 259
QY 63 ArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValThrLeuGlyValAlaAla 82
Db 260 CATCTGACCAATCCCGAGCATCTCTGGTTTCCGCTGCGGTTTCCCTGGGAGCGCGCC 319
QY 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 102
Db 320 GGAGGCTGTGAGTGCATATCTTCAACGACCTCTCCGCGCGGAGCTGAGCATCATGTTTC 379
QY 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
Db 380 TCAGCTGTGCGTTCGCGCGCGCTATGCTGCTCATGCGGGTTCGCGACGCGCTCTGGATG 439
QY 123 LeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
Db 440 CTGCTGCTCGGAAGGACGCTGACGGGCTTCGCGGGGGGCTTCACAGCTGCGCTGCATCCG 499

```



```
QY 143 valtyrileSerGluileAlaIatyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 500 GTCTAGCTGTCTGAGATGTCTCCCGCAGCGTCTTGGGGCTCTGGGGGCGCACACCCGAC 559
QY 163 LeuMetValValValGlyIleLeuLeuAlaGlyTrpValLeuGluTrpArg 182
Db 560 CTATGGCAGTGTGGATCCCTGTCCCTCTACGCCCTTGGCTCTCTGCTGCGGTGGCGC 619
QY 183 TrpLeuAlaValLeuGlyCysValProSerLeuMetLeuLeuMetCysPheMet 202
Db 620 TGGCTGGCTGTGGCGGGAGCGCTGTGCTCATCATGATCTGCTGCTCAGCTTCATG 679
QY 203 ProGluThrProArgPheLeuLeuThrGlnHiAArgGlnGlnAlaMetAlaLeu 222
Db 680 CCCAACTCGCGCGCTTCTGCTCTCTGGGGCAGGACGAGAGGCGCTTGGCGGCGCTG 739
QY 223 ArgPheLeuTrpGlySerGluGlnGly-----TrpGluaspProIleGly----- 238
Db 740 GCCTGGCTGGTGGAGCGGACGTCGATGTCCATCTGGAGTTCGAGCAGATCCAGGACAAC 799
QY 239 ---AlaGluGlnSerPheHisLeu-----AlaLeuLeuArgGlnProGlyIleTrpLys 255
Db 800 GTCCGGAGACAGACGCGGATGTCGTGGCTGAGGCAGCGGCCCCACACGCTGTCCCG 859
QY 256 PropheIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAenAlaVal 275
Db 860 CCCATCAGCGCTGTGCTGTCGCGCTCTCTGAGCAGCTGAGCGGCATCAGCGCCATC 919
QY 276 MetPheTyrAlaGluThrIlePheGluGluAlaLysPhe-----LysAspSer 291
Db 920 CTGGTCTACCTGAGTCCATCTTCGACAGCAGCGCTGCTGTCGCGGCCCGCCAGGACGAC 979
QY 292 SerLeuAlaSerValValValGlyValIleGlnValLeuPheThrAlaValAlaLeu 311
Db 980 -----GCACCATCGTGGGGCGTGGCGCTCTGTCGCTGTCGCGCGCTCT 1030
QY 312 IleMetAspArgAlaGlyArgArgLeuLeuLeuValLeuSerGlyValValMetValPhe 331
Db 1031 ACCATGACCTCGCAGCGCGCAGGTGCTCTCTGCTCAGCGGCCATCATGTTGCT 1090
QY 332 SerThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSer 351
Db 1091 GCCAACTGACTCTGGGGGTGTACATCCACTTT-----GCCCCAGCGCTCTGAGC 1141
QY 352 HisValAlaIleSerAlaProValSer-----AlaGlnProValAsp 365
Db 1142 CCCAACGACTCGGGCGCTGGAAAGCGAGTCTTGGGGGAGTCTGGCGGAGCCCTGGCA 1201
QY 366 AlaSerValGly---LeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGly 384
Db 1202 GCACCGCTGGCTACCTACCTCGGTGGCGCTGCTGCGCACCATGCTCTTCATCATGGGC 1261
QY 385 PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLeuHis 404
Db 1262 TAGCGCTGGCTGGGGTCCCATCAGCTGGCTCTCATGTCATGCTGCTGCGCCCTGGCT 1321
QY 405 ValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424
Db 1322 GCCCGTGGGTGGCTCAGGCTCTCGGTGCTGGCGAGTGGCTACCGCCCTTGGCTCT 1381
QY 425 ThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
Db 1382 ACCAAGCTCTCTGCGCAGTGGT-GAGCACCTTCGGCTCCAGGTGCTCTTCTTCTT 1440
QY 445 -SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrLysGI 464
Db 1441 CGCGGCCATCTGCTGCTGAGCGCTGTGTTCAGGCTCTGTGTCGCGGAGACCAAGGG 1500
QY 464 yLysThrLeuGluGlnIleThrAlaHisPheGlu---GlyArg 477
Db 1501 ACGGTCCTTGGAGCAGATCGAGTCTCTTTCGCGATGGGGAGA 1543
```

RESULT 9

```
US-10-342-887-1685
; Sequence 1685, Application US/10342887
; Publication No. US20040058340A1
; GENERAL INFORMATION:
; APPLICANT: Dai, Hongyue
; APPLICANT: He, Yudong
; APPLICANT: Linsley, Peter S.
; APPLICANT: Mao, Mao
; APPLICANT: Roberts, Christopher J.
; APPLICANT: Van 't Veer, Laura Johanna
; APPLICANT: Van de Vijver, Marc J.
; APPLICANT: Bernards, Rene
; TITLE OF INVENTION: Diagnosis and Prognosis of Breast Cancer Patients
; FILE REFERENCE: 9301-188-999
; CURRENT APPLICATION NUMBER: US/10/342,887
; CURRENT FILING DATE: 2003-01-15
; PRIOR APPLICATION NUMBER: 60/298,918
; PRIOR FILING DATE: 2001-06-18
; PRIOR APPLICATION NUMBER: 60/380,710
; PRIOR FILING DATE: 2002-05-14
; PRIOR APPLICATION NUMBER: 10/172,118
; PRIOR FILING DATE: 2002-06-14
; NUMBER OF SEQ ID NOS: 2699
; SEQ ID NO 1685
; LENGTH: 2487
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-342-887-1685

Alignment Scores:
Pred. No.: 9,86e-88 Length: 2487
Score: 958.50 Matches: 227
Percent Similarity: 59.22% Conservative: 78
Best Local Similarity: 44.08% Mismatches: 164
Query Match: 39.01% Indels: 47
DB: 17 Gaps: 12

US-09-886-954A-1 (1-477) x US-10-342-887-1685 (1-2487)

QY 3 ProGlu---AspProGluLutThrGlnProLeuLeuGly----- 14
Db 20 CCCGAGAGAGACCCCGCCCATGCGAGCGCTGCTGGGAGCGGAGGCGGCGGACTGACGAC 79
QY 15 -----ProProGlyGlySerAlaProArgGlyArg----- 24
Db 80 ACCTTCCCGAGAGCGCGCCCGCTGCCAGGGGACAGCGCGGGTCCGGGACCTGCGAG 139
QY 25 -----ArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPhe 42
Db 140 AAAAAAGGGTGTCTCGGCCACCTTCCGCGCAGTGTCTCGCAATTTTCAGCTTGGGTAT 199
QY 43 AlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaPro 62
Db 200 GCCCTGGGTACACATCCCTCTCATCCAGCCCTCGAGCGCTCTTGGATCCTGACCTG 259
QY 63 ArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValTrpLeuGlyAlaAla 82
Db 260 CATCTGACCAATCCAGGCATCTCTGTTTGGGTTCGTTTCCCTTGGAGGAGCGGCC 319
QY 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 102
Db 320 GGAGCGCTGAGTGCATGATCTCAACGACCTCTCCGGCGCGAAGCTGAGCATCATGTT 379
QY 103 CysSerValPropheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
Db 380 TCAGCTGTGCGCTCGCGCGCGCTATGCGCTCATGGCGGGGTGCGCAGCGCTCTGATG 439
QY 123 LeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
Db 440 CTGCTGTCTCGGAAGGACGCTGAGCGGCTTCCGCGGGGGGCTCACAGCTCCTCGATCC 499
QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
```

```

500 GTGTACGTCGTAGATTGCTCCCGAGCGGCTTGTGGGGCTCTGGGGCCACACCCAG 559
163 LeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGluTyrArg 182
560 CTCATGGCAGTGTGGATCCCTGCTAGCGCCCTGGCTCTCTGCTGGCGC 619
183 TrpLeuAlaValLeuGlyCysValProSerLeuMetLeuLeuMetCysPheMet 202
620 TGCTGGCTGTGGCGGGAGCGCTGTGTCTCATCATGATCCTGCTCAGTCTCATG 679
203 ProGluThrProArgPheLeuThrGlnHisArgGlnGlnGlnGlnGlnGlnGln 222
680 CCAACTCCCGCGCTTCTGCTCTCTGGGGCAGGAGAGAGCCCTGGGGCGCTG 739
223 ArgPheTrpGlySerGlnGlnGly-----TrpGluAspProProIleGly----- 238
740 GCCTGGCTGGGGAGCGACGTGTCATGTCCTGGAGTTCGAGCAGATCCAGGACAAC 799
239 ----AlaGlnSerPheHisLeu-----AlaLeuLeuArgGlnProGlyIleTyrIys 255
800 GTCCGAGACAGACGAGCGAGTATCGTGGCTGAGGCACGGGCCCCACACGTGTGCCG 859
256 ProPheIleGlyValSerLeuMetAlaPheGlnLeuSerGlyValAsnAlaVal 275
860 CCATCACCCTGGCTGTGTGATGGCTCTCTGAGCAGCTGACGGGATCAGCCCATC 919
276 MetPheTyrAlaGluThrIlePheGluAlaLysPhe-----LysAspSer 291
920 CTGGTCTACGTGAGTCCATCTTCGACAGCAGCGCTGCTCTGCTGCCCGCCCAAGCAG 979
292 SerLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeu 311
980 -----GCAGCCATGTTGGGGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1030
312 IleMetAspArgAlaGlyArgLeuLeuValLeuSerGlyValValMetValPhe 331
1031 ACCATGGACCTCGAGCGCGAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1090
332 SerThrSerAlaPheGlyAlaTyrPheIysLeuThrGlnGlyProGlyAsnSerSer 351
1091 GCCAACCTGACTCTGGGGCTGTACATCCACTTT-----GGCGCCAGGCTCTGAGC 1141
352 HisValAlaIleSerAlaProValSer-----AlaGlnProValAsp 365
1142 CCACAGCAGCTCGGGCGCTGAAAGCAGTCTGGGGGGAGTCTGGCGAGCCCTGCGA 1201
366 AlaSerValGly----LeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAlaGly 384
1202 GCATCCGCTGGCTACCTCACCCTGGGCGCCCTGCTGGCCACCATGCTCTCATATGGC 1261
385 PheAlaValGlyTrpGlyProIleProTrpLeuMetSerGluIlePheProLeuHis 404
1262 TAGCGCGTGGCTGGGCTCCCATCCTGCTGCTCATGCTGCTGCTGCTGCTGCTGCTGCT 1321
405 ValIysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424
1322 GCCCGTGGCGTGGCTCAGGGCTCTGGGCTGGCGAGTGGCTGAGTCCAGCGCTTGTCT 1381
425 ThrIysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLeuAla 444
1382 ACCAAGTCTTCTGCGCAGTGGT--GAGCAGCTTGGCGCTCCAGGTGCTTCTTCTTCT 1440
445 -SerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThrIysG 464
1441 CGCGGCATCTGCTGTGTGAGCTGGTGTTCACAGGCTGCTGTGTGGCGGAGCAAGGG 1500
464 YLysThrLeuGluGlnIleThrAlaHisPheGlu---GlyArg 477
1501 ACGGTCCCTGGAGCAGATCGAGTCTTCTCCGATGGGAGA 1543

; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-kB
; FILE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 109
; LENGTH: 2487
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-755-889-109

Alignment Scores:
Pred. No.: 9,86e-88
Score: 958.50
Percent Similarity: 59.22%
Best Local Similarity: 44.08%
Query Match: 39.01%
Indels: 47
Gaps: 12

US-09-886-954A-1 (1-477) x US-10-755-889-109 (1-2487)
QY 3 ProGlu---AspProGluGluThrGlnProLeuLeuGly----- 14
Db 20 CCCGAGAGACCCGGCCATCGAGGAGCGCTGTGGAGCCGAGGCGCGGACTACGAC 79
QY 15 -----ProProGlyGlySerAlaProArgGlyArg----- 24
Db 80 ACCTTCCCGAGAGCGCCCGCTCGCCAGGGGACAGGGCGGCTGGGACCCCTGGAG 139
QY 25 -----ArgValPheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPheGlyPhe 42
Db 140 AACAAAGGGGTTCCTGGCCACCTTCGCCGAGTGTTCGCAATTTTCAGCTTTGGGTAT 199
QY 43 AlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProAlaPro 62
Db 200 GCCTGTGCTACATCCCTGTCTCATCCAGCCCTGGAGCGCTCTTGGATCTCTGACCTG 259
QY 63 ArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAla 82
Db 260 CATCTGACCAATCCAGGCGCATCTGTGTTGGGTTCGCTGTTCCACCTGGGAGCAGCGCC 319
QY 83 GlyGlyValLeuGlyGlyTyrLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 102
Db 320 GGAGCCCTGAGTGCATGATCTCTCAACGACCTCTCTGGCGGAGAGTGCATCATGTTC 379
QY 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
Db 380 TCAGCTGTGCTGCGCGCGCGCTATGCTGCTCATGCGGGTGGCGGCTCTGGATG 439
QY 123 LeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
Db 440 CTGCTGCTCGAAGAGCGCTGACGGCTTCGCGGGGGGCTCACAGCTGCTGCATCCCG 499
QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 500 GTGTACGTGTGAGATTGCTCCCCAGCGCTTGTGGGGCTCTGGGGGGCCACACCCAG 559
QY 163 LeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGluTyrArg 182
Db 560 CTCATGGCAGTGTTCGATCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 619
QY 183 TrpLeuAlaValLeuGlyCysValProSerLeuMetLeuLeuMetCysPheMet 202
Db 620 TGCTGGCTGTGGCGGGAGCGGCTGCTGCTCATCATGATCCTGCTGCTGCTGCTGCTGCT 679

```



```
Db 672 |||||GGAGGTTGTTGGAAGCTTTTGGCGTTGGGAATTATCTCTTATGTGGTGCCTGTTTATATA 731
Qy 146 SerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetVal 165
Db 732 GCTGAGATTGCACCTCAAACTTGAGAGGTGGCTTGGGTGAGTGAACAGCTCTCTGTT 791
Qy 166 valValGlyLeuLeuLeuAlaTyrTrpValLeuGlyTrpValLeuGluTrpArgTrpLeuAla 185
Db 792 ACATTGGCATATTGCTGGCTTATCTGTTGGGTCTTTTGTCTCACTGGAGAGCTTGCA 851
Qy 186 valLeuGlyCysValProSerLeuMetLeuLeuMetCysPheMetProGluThr 205
Db 852 ATTCTAGGAATTTGGCTTGTACAGTATTAACTCGGATTATTTTTCATCTGATCC 911
Qy 206 ProArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAlaAlaLeuArgPheLeu 225
Db 912 CCCAGATGTTGCCAAGATGGGATGATAGATGTTTGAGACTTCTTTTGCAGTGTTA 971
Qy 226 TrpGlySerGluGlnGlyTrpGluAspProPheLeu----- 237
Db 972 CGAGGATTTGAC-----ACTGATATATCTGTTGAGTACATGAATTAAGAGATCT 1022
Qy 238 -----GlyAlaGluGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIle 253
Db 1023 GTGGCTTCAACGGGAAAAAGAGCTGCAATCCGATTTTGCAGATCTCAAGAGGAAAGATAT 1082
Qy 254 TyrLysProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValaen 273
Db 1083 TGGTCCCGTAAATGTTGGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1142
Qy 274 AlaValMetPheTyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerLeu 293
Db 1143 GGAATTTGTTCTATTCACTACCATCTTTTGCAATGTCAGGAATTTTCATCCAGCGAGCT 1202
Qy 294 AlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaLeuIleMet 313
Db 1203 GCTACAGTTGGACTTGGAGCGGTTGAGTTCAGTCACTGGAATTTCCACATGTTGGTG 1262
Qy 314 AspArgAlaGlyArgGlnLeuLeuValLeuSerGlyValValMetValPheSerThr 333
Db 1263 GACAAAGTGGCGGAGGCTGCTCTTAATAATATCTCATCTGTAATGACAGTTAGCCCT 1322
Qy 334 SerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerHisVal 353
Db 1323 CTGATTGTTCTATAGCATTTTATCTG---GAGGGGGTGTATCAGAGGATTCACATTTA 1379
Qy 354 AlaIleSerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeu 373
Db 1380 -----TTCAGCATTTTGGGAATAGTTTCTATTGTTGACTCGTGGCTATG 1424
Qy 374 AlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProIlePro 393
Db 1425 GTGATTT-----GGTTTCTCTAGTCTGGGACCCATCCCT 1460
Qy 394 TrpLeuLeuMetSerGluIlePheProLeuHisValLysGlyValAlaThrGlyIleCys 413
Db 1461 TGGCTTATAATGCTCAGATACTTCCAGTGAATATAAGGGCTTCTGGCAGCATAGCG 1520
Qy 414 ValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMetGlu 433
Db 1521 ACAATGGGAAATGGCTGATTTGCTGGGGGATCACCAGTCACTGAATCTGTTTGAAT 1580
Qy 434 ValLeuArgProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeu 453
Db 1581 ----TGGAGCAGTGGAGGACATTTACATCTACACAGTCTGAGTCCCTTACTATAGCT 1637
Qy 454 PheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAlaHis 473
Db 1638 TTTATAGCAATGTTGGTCTCTGAGACCAAGGAGACATTTGGAAGAAATTCAGTTTCC 1697
Qy 474 Phe 474
|||
```

Db 1698 TTC 1700

RESULT 13

```
US-10-425-114-26596
; Sequence 26596, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaka, Jack E.
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(5313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 26596
; LENGTH: 1797
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4371-017-F11_FLI
US-10-425-114-26596
```

Alignment Scores:

Pred. No.:	1,248-61	Length:	1797
Score:	702.00	Matches:	170
Percent Similarity:	53.56%	Conservative:	78
Best Local Similarity:	36.72%	Mismatches:	176
Query Match:	28.57%	Indels:	40
DB:	17	Gaps:	6

US-09-886-954A-1 (1-477) x US-10-425-114-26596 (1-1797)

```
Qy 27 PheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyr 46
Db 284 TTCTCTGCACGCTCATCGTGGCTCGGCCCATCCAGTTCGGATTCACGGGGGCTTC 343
Qy 47 SerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArg-LeuAspAs 66
Db 344 TCCTCCCGCAGC-----CAGGACGCCATCAT-CCGAGACCTCGACCTCACCT 390
Qy 66 palalaalaSsrTrpPheGlyAlaValThrLeuGlyAlaAlaAlaGlyGlyValle 86
Db 391 CTCGAGTTCTCGGTGTCGGATCGCTGCCAACGTCGGCGCCATGGTTGGGGCGATTGC 450
Qy 86 uGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValPr 106
Db 451 CAGTGGTCAGATGGCGGAGTACATTGGGGCGCAAGGGTTCATTGATGATTGCTGCAATTC 510
Qy 106 oPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuLeuGlyGl 126
Db 511 AAACATCATTTGGTGGCTTGGCCATCTCCCTTGCAAAAGACTCATCGCTTTCTTATATGG 570
Qy 126 yArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProValTyrIleSe 146
Db 571 ACGATTGCTCGAGGGGTTTGGTGGTGTCTATCTCTTATACGGTCCAGTTTACATAGC 630
Qy 146 rGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetValVa 166
Db 631 AGAATATACCTCAAAACATGAGAGGTGCTCTTGGCTCAGTGAATCAGTTATCTGTAA 690
Qy 166 lValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArgTrpLeuAlaVa 186
Db 691 CGTTGGTATATTGTTGGCATATTTGCTCGGCATGTTGTTGTTCTTGGAGGCTTCTTGCTGT 750
Qy 186 lLeuGlyCysValProProSerLeuMetLeuLeuLeuMetCysPheMetProGluThrPr 206
Db 751 AATAGGAATCTTGGCCTTGCACTGTGTGTGATACCTGTCCTTATTTCTTCAATCCAGAAATCC 810
```

```
QY 206 oArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAlaLeuArgPheLeuTr 226
Db 811 AAGATGGTTGGCAAGATCAATGATGATGATTTTGGAGACTTCTTACAGATTCGAG 870
QY 226 pGlySerGluGlnGlyTrpGluAspProPheGlyAlaGlu----- 240
Db 871 GGGATTTCAG-----ACTGACATCAGCGCGGAAGTGAATGATATAAGAG 915
QY 241 -----GlnSerPheHisLeuAlaLeuLeuArgGlnProG1 252
Db 916 AGCAGTAGCGTCAGCAAAACAAAGACACGATCCGTTTCAAGAAATTAACCAAGAA 975
QY 252 yIleTyrLysProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyVa 272
Db 976 ATACCGCACACCCTTAATACTAGGAATGGCTTCTGTACTGCAACAGCTAAGTGAAT 1035
QY 272 lAenAlaValMetPheTyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSe 292
Db 1036 CAATGGAATATTGTTTATGCAAGTAGCATCTTCAAGCAGCAGGTCTCACAAACAGTGA 1095
QY 292 rLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaLeuI1 312
Db 1096 CTTGGCTACATGTCATCTGGTGTATCCAGGTTCTGCTACAGAGGTTACAACTGGTT 1155
QY 312 eMetAspArgAlaGlyArgLeuLeuValLeuSerGlyValValMetValPheSe 332
Db 1156 ATTAGACAGCTGGCCGACGATCTCTTATCACTCTTCTGCTGGATGACTCTAAG 1215
QY 332 rThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGlyAsnSerSerHi 352
Db 1216 CCTCCTTCAGTGTGCTGTTGTTATTTTCTC---AAGGATAGCATTTTCAAGAGATTCTCA 1272
QY 352 sValAlaIleSerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTr 372
Db 1273 CATGTACTACAC-----TTAAGTAT 1293
QY 372 pLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyProI1 392
Db 1294 GATCTCCTTGGTGTCTTGTGGCTTTGTATTCGCTTCTCCTTCGGTATGGTGCCAT 1353
QY 392 eProTrpLeuMetSerGluIlePheProLeuHisValLysGlyValAlaThrGlyI1 412
Db 1354 TCCATGATCATAAATGTCAGAGATCTCCCGGTAGTATCAAGAGTCTCGCAGGAAGCTT 1413
QY 412 eCysValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMe 432
Db 1414 TGCAGCGCTCGCAACTGGCTTACATCTTGGAAATAACAATGACA---GCAAACTTGAT 1470
QY 432 tGluValLeuArgProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerVa 452
Db 1471 GCTTAGCTGGAGTGTGGAGGACCTTTGTCTCTACATGCTGCTGCTTTCACCT 1530
QY 452 lLeuPheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAl 472
Db 1531 CGTGTTCGTATCATCTTTGGTGGCAGACAAAGGAAGAACTCTCGAAGAGATCAATG 1590
QY 472 aHisPhe 474
Db 1591 GTCCTTC 1597
```

RESULT 14

```
US-10-437-963-61061
; Sequence 61061, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
```

```
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 61061
; LENGTH: 2003
; TYPE: DNA
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_62530C.1
US-10-437-963-61061
```

```
Alignment Scores:
Pred. No.: 1,43e-61 Length: 2003
Score: 702.00 Matches: 170
Percent Similarity: 53.56% Conservative: 78
Best Local Similarity: 36.72% Mismatches: 176
Query Match: 28.57% Indels: 40
Gaps: 6
```

US-09-886-954A-1 (1-477) x US-10-437-963-61061 (1-2003)

```
QY 27 PheLeuAlaAlaPheAlaAlaAlaLeuGlyProLeuSerPheGlyPheAlaLeuGlyTyr 46
Db 346 TTCCTCTGCAGCTCATCGTCGCGCTCGGCCCATCCAGTTCGATTCGATTCACGGGGGCTTC 405
QY 47 SerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArg-LeuAspAs 66
Db 406 TCCTCCCGCAGC-----CAGGACGCCATCAT-CCGAGACCTCGACCTCACCCCT 452
QY 66 pAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAlaGlyGlyValLe 86
Db 453 CTCAGAGTTCGCTGTCGATCGCTGTCACAGTCGCGCCCATGTTGGGGCGATGTC 512
QY 86 uGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCysSerValPr 106
Db 513 CAGTGTTCAGATGCCGAGTACATTTGGCGCAAGGGTCAATTGATGATGTTGCTGCAATTCC 572
QY 106 oPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeuGlyG1 126
Db 573 AAACATATTGTTGGCTTGCATCTCTTTGCAAAAGACTCATCTGTTCTTTATATGGG 632
QY 126 rGluIleAlaTyrProAlaValArgGlyLeuLeuGlyValAlaSerLeuValAlaProValTyrIleSe 146
Db 633 ACGATTGCTCAGGGGTTTGGTGTGTGTGTCATCTCTTATACGGTGCAGTTTACATAGC 692
QY 146 rGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMetValVa 166
Db 693 AGAAATATCACCTCAAAACATGAGAGGTGCTCTTGGCTCAGTGAATCAGTTATCTGTAAC 752
QY 166 lValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArgTrpLeuAlaVa 186
Db 753 CGTTGGTATATTTGGCATATTTGCTCGGCATGTTTGTCTTGGAGGCTTCTTGGCTGT 812
QY 186 lLeuGlyCysValProProSerLeuMetLeuMetCysPheMetProGluThrPr 206
Db 813 AATAGGAATCTGCTTGCATCTGTTGATACCTGGCTATTCTTCTTCCAGAAATCCCC 872
QY 206 oArgPheLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArgPheLeuTr 226
Db 873 AAGATGGTTGGCAAGATGAACATGATGATGATTTTGGAGACTTCTTTACAAGTCTCTGAG 932
QY 226 pGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu----- 240
Db 933 GGGATTTCAG-----ACTGACATCAGCGCGGAAGTGAATGATATAAGAG 977
QY 241 -----GlnSerPheHisLeuAlaLeuLeuArgGlnProG1 252
Db 978 AGCAGTAGCGTCAGCAAAACAAAGACACGATCCGTTTCAAGAAATTAACCAAGAA 1037
QY 252 yIleTyrLysProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyVa 272
```

```
Db 1038 ATACCCACACCCTTAATAGTAAGTGGCTTACTGTACTCAACAGCTAGTGGAAAT 1097
QY 272 lAenAlaValMetPheTyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSe 292
Db 1098 CAATGGAATATGTTTATGACAGTACATCTTCAAGCAGCAGGCTCTCAACAGTGA 1157
QY 292 rLeuAlaSerValValGlyValIleGlnValLeuPheThrAlaValAlaAlaLeu1l 312
Db 1158 CTGGCTACATGTCACCTGGTGTCTATCCAGGTTCTTGTCTACAGGATTTACAACCTGGTT 1217
QY 312 eMetAspArgAlaGlyArgAlaGlnLeuValLeuValLeuSerGlyValValMetValPheSe 332
Db 1218 ATTAGACAGAGCTGGCGCAGGATCTCTTATCATCTCTCTGCTGGGATGACTCTAAG 1277
QY 332 rThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyGlyProGlyAsnSerSerHi 352
Db 1278 CCTCCTTGCACTGGCTGTGTATTTTCTCTC---AAGGATAGCATTTTCAAGATTTCTCA 1334
QY 352 sValAlaIleSerAlaProValSerAlaGlnProValAspAlaSerValGlyLeuAlaTr 372
Db 1335 CATGTACTACACC-----TTAAGTAT 1355
QY 372 pLeuAlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValAlaGlyTrpGlyPro1l 392
Db 1356 GATCTCCTTGGTGTCTTGTGGCTTTTGTATTCGGCTTCTCTCTCGGTATGGTGCCAT 1415
QY 392 eProTrpLeuLeuMetSerGluIlePheProLeuHisValLysGlyValAlaThrGly1l 412
Db 1416 TCATGTGATCATATGTCAGAGATCTCCCGGTAGTATCAAGAGTCTCGCAGGAGCTT 1475
QY 412 eCysValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMe 432
Db 1476 TGAGACCTCGCCAACTGGCTTACATCTTGAATAACAATGACA---GCAAACTTGAT 1532
QY 432 tGluValLeuArgProTyrGlyAlaPheTrpLeuAlaSerAlaPheCysIlePheSerVa 452
Db 1533 GCTTAGCTGGAGTGTGGAGGGACCTTGTGCTCATCTGCTGAGTGTCTTCAACCT 1592
QY 452 lLeuPheThrLeuPheCysValProGluThrLysGlyLysThrLeuGluGlnIleThrAl 472
Db 1593 CGTGTTCGTATCCTTTGGTGGTCAGAGACAAAGGAGAACTCTCGAAGATACAAATG 1652
QY 472 ahisPhe 474
Db 1653 GTGCTTC 1659
```

RESULT 15

```
US-10-425-114-24968
; Sequence 24968, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; ; FILE REFERENCE: 38-21(5313)B
; ; CURRENT APPLICATION NUMBER: US/10/425,114
; ; CURRENT FILING DATE: 2003-04-28
; ; NUMBER OF SEQ ID NOS: 73128
; ; SEQ ID NO 24968
; ; LENGTH: 2049
; ; TYPE: DNA
; ; ORGANISM: Zea mays
; ; FEATURE:
; ; OTHER INFORMATION: Clone ID: LIB3689-227-G11_FLI
US-10-425-114-24968
```

Alignment Scores:

```
Pred. No.: 1.66e-61 Length: 2049
Score: 701.50 Matches: 173
Percent Similarity: 51.19% Conservative: 86
Best Local Similarity: 34.19% Mismatches: 182
Query Match: 28.55% Indels: 65
DB: 17 Gaps: 8

US-09-886-954A-1 (1-477) x US-10-425-114-24968 (1-2049)

QY 10 GlnProLeuLeuGly-----ProGlyGlySerAlaProArg 22
Db 390 AGCCGCTCATCAACACCGGAGCTGGTACCGCATGCCCGCGGGTGGCGGTATGGGC 449
QY 23 GlyArg-----ArgValPheLeuAlaAlaPheAla----- 32
Db 450 TCGCGCAATCCAGCCTCATGAGCGCTTGGGCTCCTCTCGGTTCTCTCTCGCGCAGCTC 509
QY 33 -----AlaAlaLeuGlyProLeuSerPheGlyPhe 42
Db 510 GCTATCTCGGCCACGCTCTGCACGCTTATTGTGGCGCTAGGTCCCATCCAGTTCTGGTTTC 569
QY 43 AlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaPro 62
Db 570 ACATCGGCTACTCTCGCCACG-----CAGACGCCATCATTCGCTGATCTC 617
QY 63 ArgLeuAspAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaAla 82
Db 618 GGCCTCTCCCTCTCTGAGTTCTCCCTTCCTTCGTTTCAATATTAATAGGGCGCATGGTA 677
QY 83 GlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeu 102
Db 678 GCGGCATCTCCAGTGGGCACTTGCAGATGATATCGCGCCGCAAGGGTCTCTCATGATC 737
QY 103 CysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMet 122
Db 738 GCTGGATTCCAAACATANTGGTGGCTCGCGATATCATTCGAAAGATTCCTCTTTC 797
QY 123 LeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaPro 142
Db 798 TTGTTTATGGTGGCTGCTAGAGGATTTGGAGTCGGTGTAAATATCGTATACAGTACCG 857
QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db 858 GTTTATATTGCAAAATCGCTCCTCAGGATCAGAGGGAGCTCTTGGTTCTGTCAATCAG 917
QY 163 LeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArg 182
Db 918 CTCCTCGTCACGATTGGTATATTGCTTGCCTACTTGTGGCATGTTTGTCTCGGAGA 977
QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
Db 978 ATTCCTGCTGTTCTAGGCAATTTTACCTTGTTCATCTGATTCCTGAGCTGTTCTTTG 1037
QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeu 222
Db 1038 CCTGAATCCCAAGGTGGTTGGCAAAATGGGAAAGATGGAGGATTTTGAATATTCATTG 1097
QY 223 ArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu----- 240
Db 1098 CAAGTTCTGAGAGGATTCAG-----ACAGATATCACAGAGGATAAT 1142
QY 241 -----GlnSerPheHisLeuAlaLeuLeu 248
Db 1143 GAAATAAGAGATCATTAGCATCATCGAGGAGGAGACCAACCAAGGTTCTGCTGATATC 1202
QY 249 ArgGlnProGlyIleTyrLysProPheIleIleGlyValSerLeuMetAlaPheGlnGln 268
Db 1203 AAACAGAGAGATACAGTGTCTCCCTTGTGATAGGAAATCGGTCTCTCTCTCGCAGCAG 1262
QY 269 LeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePheGluCluAlaLysPhe 288
Db 1263 CTAAAGTGGTCAATGGCATTCATTTATTCGTCGAGCATCTTCAAGAGCTGCTGTATTT 1322
```


Search completed: February 25, 2005, 23:02:57
Job time : 708 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: February 25, 2005, 18:30:22 ; Search time 3601 Seconds

(without alignments)
5042.115 Million cell updates/sec

Title: US-09-886-954A-1

Perfect score: 2457

Sequence: 1 MPEDPETQPLPGPGSSA.....CVPETKGTLEQITAFEGR 477

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DRV=xlp
-Q=/cgn2_1/USPTO_spool_p/US09886954/runat_23022005_154242_13220/app_query.fasta_1.647
-DB=EST -QPMI=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09886954 @CGN 1.1 5180 @runat_23022005_154242_13220 -NCPU=6 -ICPU=3
-NO MAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gss1:*
9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	2159	87.9	2101	3 AK081806	AK081806 Mus muscu
2	2123.5	86.4	1954	3 AF289587	AF289587 Homo sapi
3	1713	69.7	1008	9 AY414180	AY414180 Homo sapi
4	1490.5	60.7	1098	5 BX395379	BX395379 BX395379
5	1485.5	60.5	1000	9 AY414182	AY414182 Mus muscu
6	1379	56.1	1085	4 BM545247	BM545247 AGENCOURT
7	1340	54.5	932	4 BG479842	BG479842 602527373
8	1293	52.6	830	4 BG700749	BG700749 602681616
9	1269.5	51.7	991	4 B1334832	B1334832 602998959

10	1269	51.6	1004	9 AY414181	AY414181 Pan trogl
11	1267	51.6	875	6 CD557801	CD557801 AGENCOURT
12	1241.5	50.5	880	4 BM044230	BM044230 603621486
13	1238	50.4	828	4 BG478000	BG478000 602522606
14	1233	50.2	777	4 BG717034	BG717034 602689093
15	1230.5	50.1	889	4 B1757409	B1757409 603029344
16	1222.5	49.8	879	4 BG771736	BG771736 602720404
17	1193.5	48.6	838	4 BG281777	BG281777 602402292
18	1183.5	48.2	801	4 B1114765	B1114765 602861314
19	1123	45.7	671	7 CN306077	CN306077 170006000
20	1120.5	45.6	894	5 BQ919165	BQ919165 AGENCOURT
21	1119	45.5	899	4 B1916408	B1916408 603178364
22	1116	45.4	854	4 B1828656	B1828656 603078566
23	1076	43.8	813	4 B1831908	B1831908 603078776
24	1074.5	43.7	1002	4 B1828949	B1828949 603074819
25	1065.5	43.4	736	7 CK455384	CK455384 915964 MA
26	1059	43.1	833	5 B1826390	B1826390 603076257
27	1055	42.9	700	5 BQ109572	BQ109572 imagec7
28	1036.5	42.2	934	2 BE910478	BE910478 601501005
29	1030.5	41.9	773	5 BQ603775	BQ603775 MI-P-CP1-
30	990	40.3	664	2 BE986058	BE986058 UI-M-CG0P
31	980	39.9	621	2 BE986074	BE986074 UI-M-CG0P
32	973	39.6	920	2 BF308306	BF308306 601890370
33	949	38.6	2125	3 AK089246	AK089246 Mus muscu
34	948	38.6	677	4 BG707107	BG707107 602670234
35	947.5	38.6	684	4 B1907581	B1907581 603065948
36	938	38.2	735	4 BM008943	BM008943 603618748
37	933.5	38.0	2081	3 AK079650	AK079650 Mus muscu
38	931	37.9	994	5 BQ644799	BQ644799 AGENCOURT
39	912.5	37.1	683	6 BY750238	BY750238 BY750238
40	900	36.6	638	2 BB612439	BB612439 BB612439
41	889	36.2	608	4 BG079217	BG079217 H3037C03-
42	888.5	36.2	908	6 CA327128	CA327128 UI-M-FY0-
43	870	35.4	974	5 BQ944156	BQ944156 AGENCOURT
44	869	35.4	594	4 B1402088	B1402088 MI-P-CP0-
45	866	35.2	625	5 BP152931	BP152931 BP152931

ALIGNMENTS

RESULT 1	AK081806	2101 bp	mRNA	linear	HTC 03-APR-2004
LOCUS	Mus musculus	16 days embryo	head cDNA	RIKEN full-length enriched	
DEFINITION	library, clone: C130078K14 product: solute carrier family 2, (facilitated glucose transporter), member 8, full insert sequence.				
ACCESSION	AK081806				
VERSION	AK081806.1	GI:26349396			
KEYWORDS	HTC; CAP trapper.				
SOURCE	Mus musculus (house mouse)				
ORGANISM	Mus musculus				
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.				
AUTHORS	Carninci, P. and Hayashizaki, Y.				
TITLE	High-efficiency full-length cDNA cloning				
JOURNAL	Meth. Enzymol. 303, 19-44 (1999)				
MEDLINE	99279253				
PUBMED	10349636				
REFERENCE					
AUTHORS	Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.				
TITLE	Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes				
JOURNAL	Genome Res. 10 (10), 1617-1630 (2000)				
MEDLINE	20499374				
PUBMED	11042159				
REFERENCE					
AUTHORS	Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Konno, H., Akiyama, J., Nishi, K., Kitsu, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,				

Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 20530913
 PUBMED
 11076861

REFERENCE

4 The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 695-690 (2001)

JOURNAL

5 The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II team.
 Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 6 (bases 1 to 2101)

JOURNAL

REFERENCE
 AUTHORS
 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayashi, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.
 Direct Submission

TITLE

Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.
 Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
 URL: http://fantom.gsc.riken.jp/

FEATURES

source

1. .2101
 /organism="Mus musculus"
 /mol_type="cDNA"
 /strain="C57BL/6J"
 /db_xref="FANTOM DB:C130078K14"
 /db_xref="taxon:10090"
 /clone="C130078K14"
 /tissue type="head"
 /clone_lib="RIKEN full-length enriched mouse cDNA library"
 /dev_stage="16 days embryo"
 61. .1494
 /notes="unannotated protein product; putative solute carrier family 2, (facilitated glucose transporter), member 8 (MGD)|MG1:1860103, GB|NM_019488, evidence: BLASTN, 99%, match=1490"
 /codon_start=1
 /protein_id="BAC38338.1"
 /db_xref="GI:26349397"

CDS

polyA_signal
 polyA_site
 /note="putative"
 /note="putative"

ORIGIN

Alignment Scores:
 Pred. No.: 4,05e-194 Length: 2101
 Score: 2159.00 Matches: 412
 Percent Similarity: 92.47% Conservative: 30
 Best Local Similarity: 86.19% Mismatches: 34
 Query Match: 87.87% Indels: 2
 Gaps: 2
 DB: 3
 US-09-886-954A-1 (1-477) x AK081806 (1-2101)
 QY 1 MetThrProGluAspProGluGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 DB 61 ATGTCTCCCGAGGACCCCGAGGAGCGCGCTATTGGCGCCACCGGAAGCCAGGACT 120
 QY 21 ProArgGlyArgArgValPheLeuAlaAalPheAlaAlaAlaLeuGlyProLeuSerPhe 40
 DB 121 CCCGCGCGCGCGCGGCTTCTCTCGCTTCTCTCGCGCGCGCTCTGGGACCCCTCAGTTTC 180
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaTleProSerLeuGlnArgAlaAalProPro 60
 DB 181 GGCTTCGCGCTCGGCTACAGCTCCCGCGCCATCCCGAGCTTCGCGCGCGCGCGCGCG 240
 QY 61 AlaProArgLeuAspAspAlaAalAalaserTrpPheGlyAlaValValThrLeuGlyAla 80
 DB 241 GCCCTGGCGCTCGGAGACAAATGCGGCTCTCTGGTTTCGGGCGCGTCTGACCTGGCGCT 300
 QY 81 AlalaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlyLeuSerLeu 100
 DB 301 GCTGAGGGGCGATCATCTGGCGGCTGCTCTCGACCGTTCGAGGCGCGGCGAGCTCAGGCTC 360
 QY 101 LeuLeuGlySerValProPheValAlaGlyPheAlaValThrAlaAalGlnAspVal 120
 DB 361 TTGCTTCGACCGTGGCGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 420
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 DB 421 TGGATGCTGCTCGGAGGCGCGCTCTCACCGCGCTTGTCTCGGAGTGTCTGCTCATGTG 480
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
 DB 481 GCACCGGTTTACATCTCGGAAATCGCTTACCGAGTGTCCGAGGACTGCTCGGCTCCTGT 540
 QY 161 ValGlnLeuMetValValGlyLeuLeuLeuAlaTyrLeuLeuAlaGlyTrpValLeuGlu 180
 DB 541 GTGAGCTGATGTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 600
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuLeuMetCys 200
 DB 601 TGGCGCTGCTGCGCGTGTCTGCGGCTGTGTGCGCGCGCGCGCGCGCGCGCGCGCGCG 660
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAla 220
 DB 661 TACATGCCCGAGACCCCGAGTTCCTCTCACTCAACACCATGACACGAGGCGCATGCTG 720
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyValaGlu 240
 DB 721 GCCTTCGCTTCTCTGTTGGGCTCTGAGGAGGCTGGGAAGAGCCCGCTGTTGGGCTGAG 780
 QY 241 ---GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIle 259
 DB 781 CACCAAGGCTTCCAGCTGGCGCTGTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 840
 QY 260 GlyValSerLeuMetAlaPheGlnGlnLeuSerGlyValAenAlaMetPheTyrAla 279
 DB 841 GGCAATTTCCTCATGCTTCCAGGAGCTGTGAGGCGTCAATGCTATCATGTTCTATGCC 900

```

QY 280 GluThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGly 299
Db 901 AACAGCATCTTCAGGAGGCGCAAGTTCAAGGACAGCAGCTGCTCGGTCACTGGGCG 960
QY 300 ValIleGlnValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArg 319
Db 961 ATAATCAAGTCTGTTCACTGCTGTGGCGGCTCATCATGACAGCAGCGGCGAAG 1020
QY 320 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyr 339
Db 1021 CTGCTCTGGCTGTGGGTGTGATCATGCTGTGATGATGATGATGATGATGATGATGAT 1080
QY 340 PheLysLeuThrGlnGlyProGlyAanSerSerHisValAlaIleSerAlaProVal 359
Db 1081 TTCAAACTGACCCAGAGCTCCCGACCACTCTCCCGACGATAGGCTG--GTGCCCATC 1137
QY 360 SerAlaGlnProValAspAlaSerValGlyLeuAlaThrLeuAlaValGlySerMetCys 379
Db 1138 GCGCGGAGGCTGTGATGTCCAAAGTGGGACTGGCTGGCTGGCTGGCTGGCTGGCTGG 1197
QY 380 LeuPheIleAlaGlyPheAlaValGlyTyrGlyProIleProIleProIleProIlePro 399
Db 1198 CTCTTCATGCTGGCTTTCGGGTGGCTGGGTGGCTGGGTGGCTGGGTGGCTGGGTGGCT 1257
QY 400 IlePheProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeu 419
Db 1258 ATCTTCCCTCTGCATGTCAAGGTGTGGCTACCGGCTCTGTGTCTCACCACTGGTTC 1317
QY 420 MetAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGly 439
Db 1318 ATGGCTTTCTAGTGACCAAGAGTTTCAGCAGCGTTCATGAGATGTCAGACCTACGGT 1377
QY 440 AlaPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysVal 459
Db 1378 GCCTTCGTGCTACCGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTTCTGCTGCTG 1437
QY 460 ProGluThrLysGlyThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 1438 CCTGAGACTAAAGCAGGACTCTGGAAACAAGTCACAGACCCATTTTCAGGGGACGA 1491

RESULT 2
AF289587 1954 bp mRNA linear HTC 01-JAN-2002
LOCUS Homo sapiens clone pp7425 unknown mRNA.
DEFINITION AF289587
ACCESSION AF289587
VERSION AF289587.1 GI:18027377
KEYWORDS HTC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1954)
AUTHORS Zhang, P.P., Zhou, X.M., Jiang, H.Q., Huang, Y., Qin, W.X., Zhao, X.T.,
Wan, D.F. and Gu, J.R.
TITLE Novel human cDNA clones with function of inhibiting cancer cell
growth
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1954)
AUTHORS Zhang, P.P., Zhou, X.M., Jiang, H.Q., Huang, Y., Qin, W.X., Zhao, X.T.,
Wan, D.F. and Gu, J.R.
TITLE Direct Submission
JOURNAL Submitted (17-JUL-2000) National Laboratory For Oncogenes & Related
Genes, Shanghai Cancer Institute, 25/Ln 2200 Xie-Tu Road, Shanghai
200032, P. R. China
FEATURES
Location/Qualifiers
1..1954
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="pp7425"
491..1237
/codon_start=1
CDS

```

```

/product="unknown"
/protein_id="AAL55771.1"
/db_xref="GI:18027378"
/translation="VVVVGILAVLWVLEWRLAVLGVPSLSMLLMLCFMPTPTP
ELLTQRRQEAALRPFNGSQCHDEDPGAEQSFHLALLRQPKYKFPFIQVSLMA
FOQLSGVNAVFYAEIIFEEKKDSLASVVGVIQVLFATAVALIMDRAGRLILV
LSGVVVFSTSAFGAYFKLTGGPGNSHVAISAPVSAQPDVASVGLAWLAVGSMCLF
TAGGPOALWSLLACLRLFLHLCQPFHFVLCP"

```

ORIGIN

```

Alignment Scores: 8,69e-191 Length: 1954
Score: 2123.50 Matches: 426
Percent Similarity: 89.14% Conservative: 1
Best Local Similarity: 88.94% Mismatches: 1
Query Match: 86.43% Indels: 51
DB: 3 Gaps: 1

US-09-886-954A-1 (1-477) x AF289587 (1-1954)

QY 1 MetThrProGluAspProGluGlnThrProLeuLeuGlyProProGlyGlySerAla 20
Db 1 GTGACGCCGAGGACCCAGAGAAACCCAGCCGCTTCTGGGGCTTCTGGGGCGGCGG 60
QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db 61 CCCCAGCGCGCGGCTTCTTCTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCG 120
QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
Db 121 GCTTTCGCGCTCGGCTACAGCTCCCGCGCATCCCTAGCTGCGCGCGCGCGCGCG 180
QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyValAlaValThrLeuGlyAla 80
Db 181 GCCCGCGCGTGGACAGCCCGCGCTTCTGCTGGGCTTCTGCTGGGCTTCTGCTGG 240
QY 81 AlaAlaGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
Db 241 GCGCGCGGGAGTGTCTGGCGGCTGGCTGGTGGAGCCGCGCGCGCGCGCGCGCG 300
QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaGlnAspVal 120
Db 301 TTGCTGTGCTCGCTCGCTTCTGCTGGCTTTCGCGCTTTCGCGCTTTCGCGCTTTC 360
QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 361 TGGATGCTGTGGGGGGCGGCTTCTCACCGGCTTGGCTGGCTGGCTGGCTGGCTGG 420
QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaVal-ArgGlyLeuLeuGlySerC 160
Db 421 GCGCGGCTTACATCTCGAATTCGCTTACCCAGCAGTCCCGGGGCTTCTCGGCTCT 480
QY 160 sValGlnLeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuG 180
Db 481 TGTGAGCTTAATGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 540
QY 180 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetC 200
Db 541 GTGGCGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 600
QY 200 sPheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGlnAlaMetAl 220
Db 601 CTTTCATGCCGAGACCCCGGCTTCTGCTGACTCAGACAGCGCGCGCGCGCGCG 660
QY 220 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaG 240
Db 661 CGCCCTGGGCTTCTGTGTGGGCTTCCAGCAGGAGGCTGGGAAGACCCCGCGCG 720
QY 240 uGlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleG 260
Db 721 GCAGAGCTTTTACCTGGGCGCTGCTGGCGAGCGCGCGCGCGCGCGCGCGCGCG 780
QY 260 yValSerLeuMetAlaPheGlnLeuSerGlyValAlaAlaValMetPheTyrAlaG 280

```

```

Db      781  CGTCTCCCTGATGSCCTTCCAGCAGCTGTGGGGGTCAAGCCGTCATGTTCTATGCAGA 840
Qy      280  uThrIlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVa 300
Db      841  GACCATCTTTGAAGAGCCAAAGTTCAGAGCAGCAGCCTGGCTCGGTCTGGTGTGT 900
Qy      300  lIleGlnValLeuPheThrAlaValAlaLeuLeuMetAspArgAlaGlyArgGle 320
Db      901  CATCCAGGTGCTGTTCAACAGTGTGGCGGCTTCATATGACAGACAGAGCGGAGGCT 960
Qy      320  uLeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTrpH 340
Db      961  GCTCCTGGTCTGTGTCAGGTGTGTGTCAGTGTGTTTCAGCAGAGTGCCTTCGGCGCTACTT 1020
Qy      340  eLysLeuThrGlnGlyClyProGlyAsnSerSerHisValAlaLailSerAlaProValSe 360
Db      1021  CAAGCTGACCCAGCGTGGCCCTGGCAACTCTCGCACGTGGCCATCTCGCGGCTGTCTC 1080
Qy      360  rAlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLe 380
Db      1081  TGCACAGCCTGTGTATGCCAGCGTGGGGCTGGGCTGGCTGGCGGTGGCAGCATGTGCTT 1140
Qy      380  uPheIleAlaGly-PheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluI 400
Db      1141  CTTTCATCGCGGAGGTC----- 1157
Qy      400  lePheProLeuHisValLysGlyValAlaThrGlyLeCysValLeuThrAsnTrpLeuM 420
Db      1157 ----- 1157
Qy      420  etAlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTrpGlyVA 440
Db      1158 -----CTCAGGCCCTATGGAG 1173
Qy      440  laPheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValP 460
Db      1174  CCTTCCTGGCTGCTCCGCTTTCTGCATCTTCAGTGTCTTTCACCTTTGTTCTGTGTCC 1233
Qy      460  roGluThrLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db      1234  CTGAACCTAAGGAAAGACTCTGGAAACAATCACAGCCCAATTTGAGGGGCGA 1286

RESULT 3
LOCUS   AY414180
DEFINITION Homo sapiens SLC2A8 gene, VIRUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY414180
VERSION AY414180.1 GI:39770142
KEYWORDS GSS.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1008)
AUTHORS Clark,A.G., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
REFERENCE 2 (bases 1 to 1008)
AUTHORS Clark,A.G., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering

```

```

FEATURES
    source          them based on alignment.
    1..1008         Location/Qualifiers
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    <1..>1008
                    /gene="SLC2A8"
                    /locus_tag="HOM5138"
ORIGIN

```

```

Alignment Scores:
Pred. No.:      3,72e-152      Length:      1008
Score:          1713.00        Matches:    332
Percent Similarity: 99.10%      Conservative: 0
Best Local Similarity: 99.10%    Mismatches: 3
Query Match:     69.72%         Indels:     0
DB:              9             Gaps:       0

```

```

US-09-886-954A-1 (1-477) x AY414180 (1-1008)
Qy      143  ValTrpIleSerGluIleAlaTrpProAlaValArgGlyLeuLeuGlySerCysValGln 162
Db      1  GTCTACATCTCGAAATCGCTACCCAGCAGTCCGGGGGTGCTCGGCTCTCTGTGCAG 60
Qy      163  LeuMetValValValGlyIleLeuLeuAlaTrpIleAlaGlyTrpValLeuGluTrpArg 182
Db      61  CTAATGGTCGTGTCGGGCATCTCTCGCCCTACCTGGCAGCTGGGTGCTGGAGTGGCGC 120
Qy      183  TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
Db      121  TGGCTGGCTGTGCTGGGCTGGCTGCCCTCCCTCCCTCATGCTGCTTCTCATGTGCTTCATG 180
Qy      203  ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeu 222
Db      181  CCCGAGACCCCGCTCTCTGCTGACTCAGCACAGGCGCCAGAGGAGCCATGGCCGCCCTG 240
Qy      223  ArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGluGlnSer 242
Db      241  CGGTTCCTGTGGGCTCCGAGCAGGCTGGGAGAGACCCCCCATCGGGGCTGAGCAGAGC 300
Qy      243  PheHisLeuAlaLeuLeuArgGlnProGlyIleTrpLysProPheIleIleGlyValSer 262
Db      301  TTTCACTGGCCCTGCTGGCGAGCCCGGCTCTACAGGCCCTTCATCATCGGCGTCTCC 360
Qy      263  LeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTrpAlaGluThrIle 282
Db      361  CTGATGGCTTCCAGCAGCTGTGGGGGTCAACGCCCTCATGTTCTATGCAGAGACCATC 420
Qy      283  PheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValValGlyValIleGln 302
Db      421  TTTGAAGAGCCCAAGTTCAAGGACAGCAGCTGGCTGGCTGGCTGGGTGTCATCCAG 480
Qy      303  ValLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeu 322
Db      481  GTGCTGTTCAGAGCTGTGGGGCTCTCATCATGACAGACAGCAGGCGGAGGCTGCTCTG 540
Qy      323  ValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTrpPheLysLeu 342
Db      541  GTCTTNNNNNGTGTGGTTCATGGTGTTCAGCACAGAGTGCCTTCGGCGCTACTTCAAGCTG 600
Qy      343  ThrGlnGlyProGlyAsnSerSerHisValAlaLailSerAlaProValSerAlaGln 362
Db      601  ACCCAGGGTGGCCCTGGCAACTCTCTCGCAGCTGGCCATCTCGCGGCTGTCTCTGCACAG 660
Qy      363  ProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIle 382
Db      661  CCTGTTGATGCCAGCGTGGGCTGGGCTGGCTGGCTGGCTGGCAGCATGTGCTCTTTCATC 720
Qy      383  AlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhePro 402
Db      721  GCCCGCTTTCGGGTGGGCTGGGGGCCCAATCCCTTGGCTCTCTCATGTGCAGAGATCTTCCCT 780

```

QY 403 LeuHisValIysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPhe 422
 DB 781 CTGCATGTCAAGGCGTGGCGACAGGCATCTGGTCTCTACCAACTGGCTCATGGCCTTT 840
 QY 423 LeuValThrIysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrp 442
 DB 841 CTGCTGACCAAGGAGTTCAGCAGCCCTCATGGAGGTCTCTCAGGCGCTTCTGG 900
 QY 443 LeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluThr 462
 DB 901 CTTCGCTCCGCTTCTGCATCTTCAGTGCCTTTTCATTTGTCTGTCTCCCTGAACT 960
 QY 463 LysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 DB 961 AAAGGAAGACTCTGGAACAAATCACAGGCCATTTTGAGGGGCGA 1005

RESULT 4
 BX395379
 LOCUS BX395379 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
 DEFINITION CDNA clone CS0DD006YP06 5-PRIME, mRNA sequence.
 ACCESSION BX395379
 VERSION BX395379.2 GI:46875461
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 TITLE Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 JOURNAL 1 (bases 1 to 1098)
 COMMENT Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
 Full-length cDNA libraries and normalization
 Unpublished (2001)
 On May 13, 2003 this sequence version replaced gi:30620615.
 Contact: Genoscope

Genoscope - Centre National de Sequencage
 2 rue Gaston Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE
 Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
 1st strand cDNA was primed with a NotI-oligo(dT) primer. Five prime
 end enriched, double-strand cDNA was digested with Not I and cloned
 into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library
 was normalized. Library was constructed by Life Technologies, a
 division of Invitrogen. This sequence belongs to sequence cluster
 5383.f
 For more information about this cluster, see
 http://www.genoscope.cns.fr/cdna?s=CS0DD006DH03QP1kc=5383.f.
 Location/Qualifiers

FEATURES
 source
 1..1098
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="CS0DD006YP06"
 /tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
 /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
 /notes="1st strand cDNA was primed with a NotI-oligo(dT)
 primer. Five prime end enriched, double-strand cDNA was
 digested with Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized."
 sites of the pCMVSPORT 6 vector. Library was normalized."

ORIGIN
 Alignment Scores:
 Pred. No.: 6,21e-131 Length: 1098
 Score: 1490.50 Matches: 320
 Percent Similarity: 68.72% Conservative: 3
 Best Local Similarity: 68.09% Mismatches: 23
 Query Match: 60.66% Indels: 127
 DB: 5 Gaps: 1
 US-09-886-954A-1 (1-477) x BX395379 (1-1098)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
 DB 19 ATGACGCCCGAGACCCAGAGGAACCCAGCCCGCTTCTGGGGCCTCTCTGGCGGCAG---- 74

QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 DB 74 ----- 74
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 DB 74 ----- 74
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 DB 74 ----- 74
 QY 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgLysLeuSerLeu 100
 DB 74 ----- 74
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 DB 74 ----- 74
 QY 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 DB 74 ----- 74
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuGlySerCys 160
 DB 75 -----GTCTACATCTCCGAAATCGCTACCCAGCAGTWCAGGGGGTGTCTCGGCTCTGT 128
 QY 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
 DB 129 GTGACGCTAATGGTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 187
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 DB 188 TKGCGCTGGCTGGCTGT 247
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGlnAlaMetAla 220
 DB 248 TTTATGCCCGAGACCCCGGCTTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGTCTGT 307
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaGlu 240
 DB 308 GCCTCGCNGTCTCTGTGGGGCTCCGAGCAGGGCTGGGAAGACCCCGCTGTGTGTGTGTGT 367
 QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleGly 260
 DB 368 CAGAGCTTTCACCTGGCCCTGTCTGGCGAGCAGCGCATCTACAGCTCTTATCTATCGGC 427
 QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGlu 280
 DB 428 GTCTCCCTGATGGCTTCCAGCAGCTGTGGGGGTAAAGCCGCTTATGTCTATGCAGAG 487
 QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
 DB 488 ACCATCTTTGAAGAGCCCAAGTTCAGGACAGCAGCTGGCTGTGTGTGTGTGTGTGTGTGT 547
 QY 301 IleGlnValLeuPheThrAlaValAlaAlaLeuLeuMetAspArgAlaGlyArgArgLeu 320
 DB 548 ATCCAGGTGCTGTWCACAGCTGTAGCGGCTCTTATCATGACAGACAGGGCGAGGCTG 607
 QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPhe 340
 DB 608 CTCTGGTCTTGTGTAGGT 667
 QY 341 LysLeuThrGlnGlyGlyProGlyProGlyProGlyProGlyProGlyProGlyProGly 360
 DB 668 AAGCTGACCCAGGGTGGCCCTGGCAACTCTCTCGACGTGGCCATMTMGCGCGCTGTCT 727
 QY 361 AlaGlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeu 380
 DB 728 GCACAGCCTGT 787
 QY 381 PheIleAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIle 400

Db 788 TWATCGCGCGCTTGGCGTGGCGTGGCGGCGCAATCCCTGGCTCCTCATGTHGAAATC 847
 QY 401 PheProLeuHisValLysGlyValAlaThrGlyLeCysValLeuThrAsnTrpLeuMet 420
 Db 848 TTCCTCTGCATGTTAAGGGCGTGGACAGGVAATCGCGTCTCTCAACCACTGGCTCATG 907
 QY 421 AlaPheLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAla 440
 Db 908 GCCTTTCTCGTACCAAGGAGTTGACAGCTTCATGGAGGTCTCTCAGGCCCTATGAGGCC 967
 QY 441 PheTrpLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValPro 460
 Db 968 TTCTKGCTTGCTCCGCTTCTGCATCTCCAGTTT-CTTTTCATTGTTCTNTTGTCT 1026
 QY 461 -GluThrLysGlyThrLeuGluIn 469
 Db 1027 TAAACTAAAGGAAATTTTGGACCA 1054

RESULT 5
 LOCUS AY414182 1000 bp DNA linear GSS 17-DEC-2003
 DEFINITION Mus musculus SLC2A8 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.

ACCESSION AY414182
 VERSION AY414182.1 GI:39770144
 KEYWORDS GSS.
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 1000)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios

JOURNAL Science 302 (5652), 1960-1963 (2003)
 PUBMED 14671302
 REFERENCE 2 (bases 1 to 1000)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Perriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.

TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment

FEATURES
 source Location/Qualifiers
 1..1000
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"
 gene <1..>1000
 /gene="SLC2A8"
 /locus_tag="HGM5138"

ORIGIN
 Alignment Scores:
 Pred. No.: 1,62e-130 Length: 1000
 Score: 1485.50 Matches: 284
 Percent Similarity: 91.0% Conservative: 22
 Best Local Similarity: 84.5% Mismatches: 26
 Query Match: 60.4% Indels: 4
 DB: 9 Gaps: 2

US-09-886-954A-1 (1-477) x AY414182 (1-1000)

QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
 Db 1 GTTATCATCTCGGAAATCGCTACCCAGCTGTCCGAGGACTGCTCGGCTCTGTGTGACG 60

QY 163 LeuMetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTyrValLeuGluTrpArg 182
 Db 61 CTGATGTTGTCACTGGCATCTCTCTGGCTATGTGGCAGGCTGGTCTTAGAGTGGCG 120
 QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
 Db 121 TGGCTGGCGCTGTGGCTGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 180
 QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGlnAlaMetAlaLeu 222
 Db 181 CCGAGAGCCACAGTTTCTCTCTCAACACAGTACAGAGGAGCCATGGCTGCCCTG 240
 QY 223 ArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheIleGlyValAlaGlnSer 242
 Db 241 CGCTTCTCTGTGGGCTCTGAGGAGGCTGGGAGAGCCCTCTGTTGGGCTGAGCAGCG 300
 QY 243 PheHisLeuAlaLeuArgGlnProGlyIleTyrLysProPheIleGlyValSer 262
 Db 301 TTCCAGCTGGCGCTCTGAGGCGCTGGCATCTACAAGCCCTCATCATCGCATTTCC 360
 QY 263 LeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIle 282
 Db 361 CTGATGTTCTCCAGCAGCTGTCAAGGCTCAATGCTATCATGTTCTATGCCAACAGCATC 420
 QY 283 PheGluGluAlaLysPheLysAspSerSer-LeuAlaSerValValGlyValIleGly 302
 Db 421 TTCGAGGAGGCCAAGTTTCAG- - - - -CAGCTTGGCTCGGTCACTGTGGGCATAATCCA 474
 QY 302 nValLeuPheThrAlaValAlaAlaLeuLeuMetAspArgAlaGlyArgArgLeuLeu 322
 Db 475 GGTCTGTCTCACTGTGTGGCGGCTCATCATGACAGAGAGGCGGAGGCTGTCTCT 534
 QY 322 uValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLe 342
 Db 535 GGCCTTGTGGGCTGTGATCATGTGTTCAGTATCAGTGCCTTGTGTACCTACTTCAACT 594
 QY 342 uThrGlnGlyProGlyAsnSerHisValAlaIleSerAlaProValSerAlaGly 362
 Db 595 GACCAAGAGCTCCCGAGCAACTCTCCAGTAGGCGCTG- - -GTGCCCATCGCGCGGA 651
 QY 362 nProValAspAlaSerValGlyLeuAlaTyrLeuAlaValGlySerMetCysLeuPheIle 382
 Db 652 GCCTGTGATGTCCAAGTGGGACTGGCTGGCTGGCTGTAGGAGAGCATGTGCTCTTCAT 711
 QY 382 eAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhePr 402
 Db 712 TGCTGGCTTTGCGGTGGGCTGGGAGCCATCCCTGGCTCTCATGTGACAGATCTTCCC 771
 QY 402 oLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPh 422
 Db 772 TCTCATGTCAAGGGTGTGGTACCGGCATCTGTGCTCTCACCACACTGGTTCATGGCCTT 831
 QY 422 eLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTr 442
 Db 832 TCTAGTCACCAAGAGTTTCAGCAGCGTTCATGAGATGCTCAGACCTCTACGCTGCTG 891
 QY 442 pLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluTh 462
 Db 892 GCTCAGCGCTTCTGCGCTCTCAGTGTCTTATTCACACTGACCGTGTGCTCTGAGAC 951
 QY 462 rLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
 Db 952 TAAAGGCAGGACTCTGGAACAAGTACAGGCCCATTTTCGAGGGACGA 997

RESULT 6
 BM545247
 LOCUS BM545247 1085 bp mRNA linear EST 20-FEB-2002
 DEFINITION AGENCOURT 6497266 NIH_MGC_124 Homo sapiens cDNA clone IMAGE:5726945
 5', mRNA sequence.
 ACCESSION BM545247
 VERSION BM545247.1 GI:18777177
 KEYWORDS EST.


```

SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 1085)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Invitrogen
            CDNA Library Preparation: Life Technologies, Inc.
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Agencourt Bioscience Corporation
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLM12720 row: d column: 18
            High quality sequence stop: 679.
            Location/Qualifiers
              1..1085
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:5726945"
                /tissue_type="hippocampus"
                /lab_host="DH10B"
                /clone_lib="NIH MGC 124"
                /note="Organ: Brain; Vector: pCMV-SPORT6; Site 1: EcorV
                (destroyed); Site 2: NotI; RNA source male hippocampus,
                age 27. Library is oligo-dT primed and directionally
                cloned (Scorv site is destroyed upon cloning). Average
                insert size 1.4 kb, insert size range 0.9-4 kb. Library is
                normalized and enriched for full-length clones and was
                constructed by C. Gruber (Invitrogen). Research Genetics
                tracking code 012."

ORIGIN
Alignment Scores:
Pred. No.:      2,47e-120      Length:      1085
Score:          1379.00      Matches:     279
Percent Similarity: 87.77%      Conservative: 1
Best Local Similarity: 87.46%      Mismatches: 6
Query Match:     56.13%      Indels:      33
DB:              4              Gaps:         3

US-09-886-954A-1 (1-477) x BME45247 (1-1085)

QY      162 GlnLeuMetValValGlyLeuLeuAlaTyrrLeuAlaGlyTrpValLeuGluTrp 181
Db      1 CAGCTAATGGTCGTCTGGCATCTCTCTGGCTACCTGGCAGGCTGGGTGGAGTGG 60

QY      182 ArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPhe 201
Db      61 CGCTGGCTGGCTGTGTGGGCTGGTGGCCCTCCCTCATGCTGTCTCATGTGCTTC 120

QY      202 MetProGluThrProArgPheLeuLeuThrGlnHisArgGlnGluAlaMetAlaAla 221
Db      121 ATGCCCGAGACCCCGCGCTTCCTGCTGACTCAGCAGCGGCGCAGAGCCCATCGGGCTGAGCAG 180

QY      222 LeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProLeuGlyAlaGluGln 241
Db      181 CTGCGGTTCCTGTGGGGCTCCGAGCAGGGCTGGGAAGACCCCCCATCGGGCTGAGCAG 240

QY      242 SerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrrLysProPheIleLeuVal 261
Db      241 AGCTTTTACCTGGCCCTGTGGGCGACCGCGCATCTACAAGCCCTTCATCATCGCGCTC 300

QY      262 SerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrrAlaGluThr 281
Db      301 TCCTGATGGCTTCAGCAGCTGTCTGGGGGTCAACGCCGTATGTTCTATCAGAGACC 360

QY      282 IlePheGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyValIle 301

```

```

Db      361 ATCTTTGAAGAGCCCAAGTTCAAGGACAGCAGCTGGCTCGGTCTGCTGGGTGTCATC 420
QY      302 GlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgGluLeu 321
Db      421 CAGGTGCTGTTTCACAGCTGTGGCGCTCTCATATGACAGACAGCGGGGAGGCTGCTC 480
QY      322 LeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrrPheLys 341
Db      481 CTGGTCTTGTCAAGTGTGGTGTGTTTCAGCAGCAGAGTGCCTTCGGCGCCTACTTCAAG 540
QY      342 LeuThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAla 361
Db      541 CTGACCCAGGGTGGCCCTGGCAACTCTCGCACGTGGCCATCTCGGCGCTGTCTCTGCA 600
QY      362 GlnProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPhe 381
Db      601 CAGCCTGTGTATGCCAGCGTGGCGCTGGCTGGCTGGCGGAGCATGTGGCTCTTC 660
QY      382 IleAlaGlyPheAlaVal-GlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePh 401
Db      661 ATCGCGGCTTTTGGGCTGGGGGCCCATCCCTGGCTCTCATGTCTCAGAGATCTT 720
QY      401 eProLeuHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAl 421
Db      721 CCTCTGTCATGTCAAGGGCGTGGCGACAGGCATCTGGCTCTCACCACTGGCTCATGGC 780
QY      421 apheLeuValThrLysGlyPheSerSerLeuMet----- 432
Db      781 CTTTCTNTGACCAAGGAGTTCAGCAGCCTCATGTCTCTGGCTGCTCTGAGGACTCAN 840
QY      433 -----GluValLeuArgPro----- 437
Db      841 GAACACCTTCAGACTTTCAGACCTTCGGGTGAGCCCTCCATCGGCAAGATAAAGCAACGG 900
QY      438 -----TyrrGly-----AlaPheTrpLeuAlaSerAlaPhe 447
Db      901 NAAAAAGGAGGGGGCGCTCAAGGACCTTTGGCTTTCTGGCTTGAAGGGCCTTTT 955

RESULT 7
BG479842      602527373F1 NIH_MGC_21 Homo sapiens cDNA clone IMAGE:4650906 5',
LOCUS      mRNA sequence.
DEFINITION      BG479842
ACCESSION      BG479842
VERSION      BG479842.1 GI:13412121
KEYWORDS      EST.
SOURCE      Homo sapiens (human)
ORGANISM      Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 932)
AUTHORS    NIH-MGC http://mgc.nci.nih.gov/.
TITLE      National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL    Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: ATCC
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            CDNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: LLM1433 row: a column: 19
            High quality sequence stop: 820.
            Location/Qualifiers
              1..932
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="IMAGE:4650906"
                /tissue_type="choriocarcinoma"
            source

```



```
Db      5  GCTGGGGTGTTCCTCCCTAGTGGCCCGGTATCATCTCCGAATCGCCTACCCAGCA 64
Qy      153 ValArgGlyLeuLeuGlySerCysValGlnLeuMetValValValGlyIleLeuLeuAla 172
Db      65  GTCCGGGGTGTCTGGCTCTCTGTGTGACAGTAAATGGTGTGTGGCATCTCTCTGGCC 124
Qy      173 TyrLeuAlaGlyTrpValLeuGluTrpArgTrpLeuAlaValLeuGlyCysValPro 192
Db      125 TACCTGGCAGGCTGGGTGTGGAGTGGCGTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 184
Qy      193 SerLeuMetLeuLeuMetCysPheMetProGluThrProArgPheLeuLeuThrGln 212
Db      185 TCCTCATGTCTCTCATGTCTCTCATGCCGAGACCCCGGCTCTCTGTGTGACTCAG 244
Qy      213 HisArgGlnGlnAlaMetAlaLeuAtrGpPheLeuTrpGlySerGluGlnGlyTrp 232
Db      245 CACAGCGCCAGAGGCCATGGCGGCTGGGTCTCTGTGGGCTCCAGCAGGCTGG 304
Qy      233 GluAspProProIleGlyAlaGluGlnSerPheHisLeuAlaLeuLeuArgGlnProGly 252
Db      305 GAAGACCCCCATCGGGCTGAGCAGAGCTTTCACCTGGCCCTGTGGCGAGCCGGC 364
Qy      253 IleTyrIysProPheIleIleGlyValSerLeuMetAlaPheGlnGlnLeuSerGlyVal 272
Db      365 ATCTCAAGCCCTTCATCATCGCGCTCTCCCTGATGGCTTCAGCAGCTGTGGGGTTC 424
Qy      273 AsnAlaValMetPheTyrAlaGluThrIlePheGluGluAlaLysPheLysAspSerSer 292
Db      425 AACGCCGTATGTTCTATGACGAGACATCTTTGAGAGGCCAAGTTTCAAGCAGCAGC 484
Qy      293 LeuAlaSerValValValGlyValIleGlnValLeuPheThrAlaValAlaLeuIle 312
Db      485 CTGGCCCTCGCTCGCTGGGTGTCTATCCAGGTCTGTTCACAGCTGTGGCGCTCTCATC 544
Qy      313 MetAspArgAlaGlyArgArgLeuLeuValLeuLeuSerGlyValValMetValPheSer 332
Db      545 ATGGACAGAGCAGCGCGAGGCTGTCTCTGGTCTGTCTAGGTGTGTGTGTGTGTGTGT 604
Qy      333 ThrSerAlaPheGlyAlaTyrPheLysLeuThrGlnGlyProGlyValSerSerHis 352
Db      605 ACAGTGGCTTTCGGCCCTACTTCAAGCTGACCCAGGTGGCCCTGGCACTCTCGCAC 664
Qy      353 ValAlaIleSerAlaProValSerAla-GlnProVal-AspAlaSerValGlyLeuAla 372
Db      665 GTGGCCATCTCGCGCTGTCTCTGCTGCACAGCTGTTTGTATGCCAGCGTGGGCTGGCT 724
Qy      372 rpleu-AlaValGlySerMetCysLeuPheIleAlaGlyPheAlaValGlyTrpGlyPro 391
Db      725 GGCTGGGCGGTGGGAGCATGTGCTCTTTCATCGCGCTTGTGGGTGGGCTGGGCGCC 784
Qy      392 IleProTrpLeuLeuMetSerGluIlePheProLeuHisVal 405
Db      785 ATCCCTTCTCTCTCTATGACAGAGATTTTCCCTCTTGTGTGTCT 826

RESULT 9
LOCUS   BI334832
DEFINITION 991 bp mRNA linear EST 30-JUL-2001
602998959F1 NIH_MGC_12 Homo sapiens cDNA clone IMAGE:5141007 5',
mRNA sequence.
ACCESSION BI334832
VERSION   BI334832.1 GI:15019489
KEYWORDS EST.
SOURCE   Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 991)
NIH-MGC http://mgi.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs@mail.nih.gov
```

Tissue Procurement: ATCC
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: Incyte Genomics, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
DNA Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
http://image.llnl.gov
Plate: LLM11347 row: f column: 16
High quality sequence stop: 835.
Location/Qualifiers
1..991
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:5141007"
/tissue_type="cervical carcinoma cell line"
/lab_host="DH10B"
/clone_lib="NIH_MGC_12"
/note="Organ: cervix; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.4 kb. Library prepared by Life
Technologies."

FEATURES

Source

ORIGIN

Alignment Scores:

Pred. No.: 5,668-110 Length: 991
Score: 1269.50 Matches: 290
Percent Similarity: 65.05% Conservative: 6
Best Local Similarity: 63.74% Mismatches: 21
Query Match: 51.67% Indels: 140
DB: 4 Gaps: 6

US-09-886-954A-1 (1-477) x BI334832 (1-991)

Qy 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAla 20
Db 17 ATGACGCCGAGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGCGAG---- 72
Qy 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
Db 72 ----- 72
Qy 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaProPro 60
Db 72 ----- 72
Qy 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
Db 72 ----- 72
Qy 81 AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspAlaGlyArgLysLeuSerLeu 100
Db 72 ----- 72
Qy 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
Db 72 ----- 72
Qy 121 TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
Db 72 ----- 72
Qy 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
Db 73 -----GCTACATCTCCGAATCGCTACCCAGCAGTCCGGGGGTGTCTCGGCTCTCT 126
Qy 161 ValGlnLeuMetValValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGlu 180
Db 127 GTGCAGCTAAATGTCGTGTGGCATCTCTCTGGCTACCTCTGGCGCTACCTGGCAGGCTGGTGGAG 186
Qy 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
Db 187 TGGCGCTGGCTGGCTGTGTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCTGGCT 246

QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 Db 247 TTCATGCCCGAGACCCCGGCTTCTGCTGACTCAGCAGCGCCGAGGAGCCATGATGCC 306
 QY 221 AlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyValaGlu 240
 Db 307 GCCCTGGGCTTCTGCGGCTCCAGCAGGCTGGAGAGCCCCCATCGGGCTGAG 366
 QY 241 GlnSerPheHisLeuAlaLeuLeuArgGlnProGlyIleTyrlYsPhePheIleIleGly 260
 Db 367 CAGAGCTTTCACGTGCGCTGCGGCGAGCCCGGCTCATCAAGCCCTTTCATCGGC 426
 QY 261 ValSerLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrlaGlu 280
 Db 427 GTCTCCCTGATGCGCTTCCAGCAGCTGTCGGGGGTCAACGCCGTGATGTTCTATGCAG 486
 QY 281 ThrIlePheGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyVal 300
 Db 487 ACCATCTTTGAGAGGCCAAGTTCAAGGACAGCAGCTGGCTCGTGGTGGTGTG 546
 QY 301 IleGlnValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArgLeu 320
 Db 547 ATCCAGGTGCTTTCACAGCTGTCGGGCTCTCATCATGACAGCAGCGCGAGGCTG 606
 QY 321 LeuLeuValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrlPhe 340
 Db 607 CTCCTGGCTTGTGAGGTGCTGATGTCAGCGTGGGCTGGCTGGCTGGCTGGCTGGCTGG 666
 QY 341 LysLeuThrGlnGlyProGlyAenSerSerHisValAlaIleSerAlaProValise 360
 Db 667 AGCTGACCCAGGCTGGCTGGCACTCTCGCAACGTGGCCATCTCGGCGCTGTCTC 726
 QY 360 rAlaGln-ProValAspAlaSerVal-GlyLeuAlaTrpLeuAlaVal-GlySerMetCy 379
 Db 727 TGACAAAGCTGCTGTGATGTCAGCGTGGGCTGGCTGGCTGGCTGGCTGGCTGGCTGG 786
 QY 379 sLeuPheIleAlaGlyPheAlaValGly---TrpGlyProIleProTrpLeu---LeuMe 397
 Db 787 CTCCTTCATGCGCGCTTTCGGTGGGCTGGGCGGCCCATCTCCCTGGTCCCTCATTTG 846
 QY 397 tSerGluIlePheProLeuHisValLysGlyValAlaThrGly-----II 412
 Db 847 TCAGAAGATCTTTCCTC-----TGCAATGTGTCAGGCGCGTGGCGCAGGCGCAT 897
 QY 412 eCyValLeuThrAsnTrpLeuMetAlaPheLeuValThrLysGluPheSerSerLeuMe 432
 Db 898 CTAGGTCCCAAAATGGCTCCATGGCTTTCGCGGACCCAG-GAGTTCAGCCAGCTCTT 956
 QY 432 tGluValLeuArgProTyrlGlyAlaPheTrpLeuAla 444
 Db 957 GGAAGTCCCAAGCCCTTGAC---CTCTGGGTGCT 990
 RESULT 10
 LOCUS AY414181
 DEFINITION Pan troglodytes SLC2A8 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY414181
 VERSION AY414181.1 GI:39770143
 KEYWORDS GSS.
 SOURCE Pan troglodytes (chimpanzee)
 ORGANISM Pan troglodytes
 Sukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
 REFERENCE 1 (bases 1 to 1004)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)

14671302
 2 (bases 1 to 1004)
 REFERENCE Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
 AUTHORS Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 TITLE Direct Submission
 JOURNAL Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1..1004
 /organism="Pan troglodytes"
 /mol_type="genomic DNA"
 /db_xref="taxon:9598"
 <1..>1004
 /gene="SLC2A8"
 /locus_tag="HCM5138"
 ORIGIN
 Alignment Scores:
 Pred. No.: 6.43e-110 Length: 1004
 Score: 1269.00 Matches: 263
 Percent Similarity: 79.88% Conservative: 3
 Best Local Similarity: 78.98% Mismatches: 67
 Query Match: 51.65% Indels: 0
 DB: 9 Gaps: 0
 US-09-886-954A-1 (1-477) x AY414181 (1-1004)
 QY 145 IleSerGluIleAlaTyrlProAlaValArgGlyLeuLeuGlySerCysValGlnLeuMet 164
 Db 3 ATCTCCGAATCGCTACCCAGCAGTCCGGGGTGTCTGGCTCTTGTGTGAGCTGATG 62
 QY 165 ValValValGlyIleLeuLeuAlaTyrlLeuAlaGlyTrpValLeuGluTrpArgTrpLeu 184
 Db 63 GTCTGCTGCGCATCTCTGCGCTTACCTGGCAGNNNNNNNNNNNNNNNNNNNNNNNN 122
 QY 185 AlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMetProGlu 204
 Db 123 NNN 182
 QY 205 ThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaLeuArgPhe 224
 Db 183 NNN 242
 QY 225 LeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGluGlnSerPheHis 244
 Db 243 NNN 302
 QY 245 LeuAlaLeuLeuArgGlnProGlyIleTyrlYsPheIleIleGlyValSerLeuMet 264
 Db 303 CTGGCCCTGCTGCGGAGCCCGCATCTACAGGCCCTTCGTCATCGGCTCTCCCTGATG 362
 QY 265 AlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrlaGluThrIlePheGlu 284
 Db 363 GCCITTCAGCAGCTGTCGGGGGTCAACGCCGTCTATGTCAGAGACCATCTTTGAA 422
 QY 285 GluAlaLysPheLysAspSerSerLeuAlaSerValValGlyValIleGlnValLeu 304
 Db 423 GAGGCCAAAGTTCAAGACAGCAGCTGGCTCGGCTCGGTCATCGGGGTGTCTCAGGTGCTG 482
 QY 305 PheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuValLeu 324
 Db 483 TTCACAGCTGTGGCGCTCTCATCATGACAGCAGCGCGGAGGCTGCTCTCTGCTCNG 542
 QY 325 SerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrlPheLysLeuThrGln 344
 Db 543 TCAGGTGGTGTGCTGCTTTCAGCAGCAGTGCCTTCGGCGCTACTTCAAGCTGACCCAG 602
 QY 345 GlyGlyProGlyAenSerSerHisValAlaIleSerAlaProValSerAlaGlnProVal 364

Db 603 GGTGGCCCTGGCAACTCTCTCGCAGTGGGCCCTCTCGCGCGCTGTCTCTGACAGAGCTGTT 662
Qy 365 AspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheLeuAlaGly 384
Db 663 GATGCCAGCGTGGGGCTGGCTGGCTGGCGTGGCGTGGCGAGCATGTCTCTTCATCGCGGC 722
Qy 385 PheAlaValGlyTrpGlyProPheLeuLeuMetSerGluLeuPheProLeuHis 404
Db 723 TTTCGGGTGGGTGGGGGCCCATCTCCCTGCTCTCATGTGACAGAGATCTTCCCTCTGTCAT 782
Qy 405 ValLeuGlyValAlaThrGlyLeuCysValLeuThrAsnTrpLeuMetAlaPheLeuVal 424
Db 783 GTCAAGGGGTGGCGAGCAGCATCTCGCTCTCACCACACTGGCTCAAGCCCTTCTCGTG 842
Qy 425 ThrLeuGluPheSerSerLeuMetGluValLeuArgProTrpGlyAlaPheTrpLeuAla 444
Db 843 ACCAAGAGTTTCAGCAGCCTCATGGAGTCTCTCAGGCCCTATGGAGCCCTTCTGGCTTGC 902
Qy 445 SerAlaPheCysLeuPheSerValLeuPheThrLeuPheCysValProGluThrLeuGly 464
Db 903 TCCGCTTCTGTCATCTTCAGTGTCTCTTTCACCTTCTGCTGTCTGCTGCTGAACTAAAGGA 962
Qy 465 LysThrLeuGluGlnLeuThrAlaHisPheGluGlyArg 477
Db 963 AAGACTCTGAAACAATCACAGCCCATTTTGAGGGCGCA 1001

RESULT 11

CD557801 875 bp mRNA linear EST 11-JUN-2003
LOCUS AGENCOURT_14413291 NIH_MGC_180 Homo sapiens cDNA clone
DEFINITION IMAGE:30386912 5', mRNA sequence.

ACCESSION

CD557801

VERSION

CD557801.1 GI:31583869

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

REFERENCE

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 875)

NIH-MGC http://mgs.nci.nih.gov/.

National Institutes of Health, Mammalian Gene Collection (MGC)

Unpublished (1999)

Contact: Daniela S. Gerhard, Ph.D.

Office of Cancer Genomics

National Cancer Institute / NIH

Bldg. 31 Rm10A07 Bethesda, MD 20892

Email: cgabs-r@mail.nih.gov

Tissue Procurement: Dr. Michael Brownstein

cDNA Library Preparation: Invitrogen Corp

cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)

DNA Sequencing by: Agencourt Bioscience Corporation

Clone distribution: MGC clone distribution information can be

found through the I.M.A.G.E. Consortium/LLNL at:

http://image.llnl.gov

Plate: NDAM456 row: f column: 09

High quality sequence stop: 618.

FEATURES

Location/Qualifiers

1..875

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="IMAGE:30386912"

/lab_host="DH10B-Ton A (T1 and T5 phage resistances) "

/clone_lib="NIH_MGC_180"

/notes="Organ: Testis; Vector: pCMV-SPORT6.1; Site 1: NotI;

Site 2: EcoRV (destroyed); Library is oligo-dT primed and

directionally cloned (EcoRV site is destroyed upon

cloning). Average insert size 1.68 kb. Library was

constructed by (Invitrogen). Note: this is a NIH_MGC

Library."

ORIGIN

Alignment Scores:

SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 880)
 AUTHORS NIH-MGC http://mgi.nci.nih.gov/
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue procurement: DCTD/DTF
 cDNA Library Preparation: Ling Hong/Rubin Laboratory
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: LLCM1930 row: k column: 22
 High quality sequence stop: 869.
 FEATURES
 Location/Qualifiers
 1..880
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:5447181"
 /tissue_type="carcinoma, cell line"
 /lab_host="DH10B (phage-resistant)"
 /clone_lib="NIH_MGC_40"
 /note="Organ: prostate; Vector: pOTB7; Site 1: XhoI;
 Site 2: EcoRI; cDNA made by oligo-dT priming.
 Directionally cloned into EcoRI/XhoI sites using the
 following 5' adaptor: GGCACGAG(G). Library constructed by
 Ling Hong in the laboratory of Gerald M. Rubin (University
 of California, Berkeley) using ZAP-cDNA synthesis kit
 (Stratagene) and Superscript II RT (Life Technologies).
 Note: this is a NIH_MGC Library."
 ORIGIN
 Alignment Scores:
 Pred. No.: 2,19e-107 Length: 880
 Score: 1241.50 Matches: 280
 Percent Similarity: 58.82% Conservative: 0
 Best Local Similarity: 58.82% Mismatches: 5
 Query Match: 50.53% Indels: 193
 DB: 4 Gaps: 2
 US-09-886-954A-1 (1-477) x BM044230 (1-880)
 QY 4 GluAspProGluThrGlnProLeuLeuGlyProProGlyGlySerAlaProArgGly 23
 Db 4 GAGGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGAG----- 50
 QY 24 ArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPheGlyPheAla 43
 Db 50 ----- 50
 QY 44 LeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProProAlaProArg 63
 Db 50 ----- 50
 QY 64 LeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAlaAlaGly 83
 Db 50 ----- 50
 QY 84 GlyValLeuGlyTyrTrpLeuValAspArgAlaGlyArgLysLeuSerLeuLeuLeuCys 103
 Db 50 ----- 50
 QY 104 SerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspValTrpMetLeu 123
 Db 50 ----- 50
 QY 124 LeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuValAlaProVal 143

Db 51 -----GTC 53
 QY 144 TyrIleSerGluLeuAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGlnLeu 163
 Db 54 TACATCTCCGAAATCGCTACCCAGCAGATCCGGGGTGTCTGGCTCTCTGTGTGACGCTA 113
 QY 164 MetValValGlyIleLeuLeuAlaTyrLeuAlaGlyTrpValLeuGluTrpArgTrp 183
 Db 114 ATGGTGTGTGTGGCATCTCTCTGGCTACTGCA----- 149
 QY 184 LeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuLeuMetCysPheMetPro 203
 Db 149 ----- 149
 QY 204 GluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaAlaLeuArg 223
 Db 149 ----- 149
 QY 224 PheLeuTrpGlySerGluGlnGlyTrpGluAspProIleGlyAlaGluGlnSer-Ph 243
 Db 150 -----GAGCTT 155
 QY 243 eHisLeuAlaLeuLeuArgGlnProGlyIleTyrLysProPheIleIleGlyValSerLe 263
 Db 156 TCACCTGGCCCTGTCTGGCGCAGCCCGGCATCTACAGCCCTTCATCATCGCGCTCCCT 215
 QY 263 uMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyrAlaGluThrIlePh 283
 Db 216 GATGGCTTCAGCAGCTGTGGGGGTCAACGCCGTGATGTTCTATGACAGACCATCTT 275
 QY 283 eGluGluAlaLysPheLysAspSerSerLeuAlaSerValValGlyValIleGlnVa 303
 Db 276 TGAAGAGCCCAAGTTCAAGGACAGCAGCTGGCCTGGTCTGTGGGTGTATCCAGGT 335
 QY 303 lLeuPheThrAlaValAlaAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuVa 323
 Db 336 GCTGTTCACAGCTGTGGCGCTCTCATCATGACAGCAGCGCGGAGGTGTCTCTGGT 395
 QY 323 lLeuSerGlyValValMetValPheSerThrSerAlaPheGlyAlaTyrPheLysLeuTh 343
 Db 396 CTTGTCAAGT 455
 QY 343 rGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaGlnPr 363
 Db 456 CCAGGTGGCCCTGCAACTCTCGACGTGGCCATCTCGCGCTGTCTCTGACAGCC 515
 QY 363 oValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIleAl 383
 Db 516 TGTGTATGCCAGCGTGGGGTGGCTGGCTGGCGGTGGCGAGCATGTGCTCTTTCATCGC 575
 QY 383 aGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePheProLe 403
 Db 576 CGCTTTTGGGTGGGTGGGGGCCCTCCCTCTCTCATGTTCAGAGATCTTCCCTCT 635
 QY 403 uHisValLysGlyValAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPheLe 423
 Db 636 GCATGTCAAGGGCGTGGCGACAGGCATCTGGCTCTCTCACCAGCTGGCTCATGGCTTTCT 695
 QY 423 uValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrGlyAlaPheTrpLe 443
 Db 696 CGTCAAGAGAGTTCAGCAGCCCTCATGGAGGTCTCAGGCCCTATGAGCCCTTCTGGCT 755
 QY 443 uAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCys-ValProGluThrL 463
 Db 756 TGCTTCCTGTTCTGCATCTTCAGTGTCTTTTTCATTTTGTGTGTGTGTGTGTGTGTGT 815
 QY 463 yGlyLysThrLeuGlnGlnIleThrAlaHisPheGluGlyArg 477
 Db 816 AAGGAAAGACTCTGGAAACAAATCACAGCCA---TTTGAGGGGGCGA 857
 RESULT 13
 BG478000

```
LOCUS      BG478000      828 bp      mRNA      linear      EST 21-MAR-2001
DEFINITION 602522606F1 NIH_MGC_20 Homo sapiens cDNA clone IMAGE:4641145 5',
mRNA sequence.
ACCESSION  BG478000
VERSION     BG478000.1  GI:13410279
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 828)
AUTHORS     NIH-MGC http://mgc.nci.nih.gov/.
TITLE       National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL     Unpublished (1999)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-x@mail.nih.gov
            Tissue Procurement: ATCC/DCTD/DRP
            CDNA Library Preparation: Ling Hong/Rubin Laboratory
            CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
            DNA Sequencing by: Incyte Genomics, Inc.
            Clone distribution: MGC clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            http://image.llnl.gov
            Plate: L1C41407 row: k column: 02
            High quality sequence stop: 788.

FEATURES             Location/Qualifiers
     source           1..828
                     /organism="Homo sapiens"
                     /mol_type="mRNA"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:4641145"
                     /issue_type="melanotic melanoma"
                     /lab_host="DH10B (phage-resistant)"
                     /clone_lib="NIH_MGC_20"
     note="Organ: skin; Vector: pOTB7; Site 1: XhoI; Site 2:
     EcoRI; cDNA made by oligo-dT priming. Directionally
     cloned into EcoRI/XhoI sites using the following 5'
     adaptor: GGACGAG(G). Size-selected >500bp for average
     insert size 1.8kb. Library constructed by Ling Hong in
     the laboratory of Gerald M. Rubin (University of
     California, Berkeley) using ZAP-cDNA synthesis kit
     (Stratagene) and Superscript II RT (Life Technologies)."
```

Alignment Scores:

Pred. No.:	4,32e-107	Length:	828
Score:	1238.00	Matches:	250
Percent Similarity:	95.08%	Conservative:	1
Best Local Similarity:	94.70%	Mismatches:	8
Query Match:	50.39%	Indels:	5
DB:	4	Gaps:	0

US-09-886-954A-1 (1-477) x BG478000 (1-828)

```
QY      1  MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
DB      47  ATGACGCCCGAGACCCAGAGAAACCCAGCGCTTCTGGGGCTCTCTGGCGGCGACGCG 106
QY      21  ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
DB      107  CCCCAGCGCGCGCGCTTCTTCGCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTTCGCGCTTC 166
QY      41  GlyPheAlaLeuGlyTyrSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
DB      167  GCGTTCCGCTCGGCTACAGCTCCCGCGCAICCTTAGCTGACGCGCGCGCGCGCGCGCG 226
QY      61  AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
DB      227  GCCCGCGCGCTGCACGACGCGCGCGCTCTCTGTTGGGGCTGCTCGTGACCTCGGGTGC 286
QY      81  AlaAlaGlyGlyValLeuGlyGlyTrpLeuValAspAlaGlyArgLysLeuSerLeu 100
DB      287  GCGGCGGGGGAGTGTGGCGGCTGGCTGGTGGTGGACCGCGCGCGCGCGCAAGCTGAGCCTC 346
```

```
QY      101  LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
DB      347  TTGCTGTGCTCGTGGCTTTCGTCGCGCTTTCGCTCATCACCGCGCGCGCGCGAGCTG 406
QY      121  TrpMetLeuLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
DB      407  TGGAGTGTGCTGGGGGCGCGCTCTCTCACCGGCTGGCGCTGGCGTGTTCCTCCCTAGTG 466
QY      141  AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCys 160
DB      467  GCCCGGCTCTCATCTCCGAAATCGCTACCCAGCAGTCCGGGGGTGTCTCGGCTCCTGT 526
QY      161  ValGlnLeuMetValValValGlyIleLeuLeuAla--TyrLeuAlaGlyTrpValLeuG 180
DB      527  GTGAGCTAATGTCGTGTCGTCGCTCTCTTGGGCTTACCTGGCAGGCTGGGTGCTGCG 586
QY      180  LuTrpArgTrpLeuAlaValLeuGlyCysVal--ProProSerLeuMetLeuLeuMet 199
DB      587  AGTGGCGCTGGCTGGCTGTGCTGGGCTGGCTTTCGCGCTTTCCTCATGTGCTTCTCATG 646
QY      200  CysPheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMet 219
DB      647  TGTTCATGCCCCGAGACCCGAGCTTCTCTGCTGACTCAGCAGGCGCCGAGGAGCCATG 706
QY      220  AlaAlaLeuArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProProIleGlyAla 239
DB      707  GCGCGCTGGGTTCTCTGTGAGGCTCCGAGCAGGCGTGGGAGAGCCCCCATCGGGCT 766
QY      240  GluGlnSerPheHisLeuAla--LeuLeuArgGlnProGlyIleTyrLysProPheIleI 259
DB      767  GAGCAGAGCTTTTACCTTGGCCCTTGTTCGGAAGCCCGGCGCATCTACAAGCCCTTCATCA 826
QY      259  le 259
DB      827  TC 828
```

RESULT 14
BG717034
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT

777 bp mRNA linear EST 08-MAY-2001
602689093F1 NIH_MGC_97 Homo sapiens cDNA clone IMAGE:4821598 5',
mRNA sequence.
BG717034
BG717034.1 GI:13996221
EST.
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 777)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-x@mail.nih.gov
Tissue Procurement: Miklos Palkovits, M.D., Ph.D.
cDNA Library Preparation: Michael J. Brownstein (NHGRI), Shiraki
Toshiyuki and Piero Carninci (RIKEN)
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM10728 row: i column: 23
High quality sequence stop: 722.
Location/Qualifiers
1..777
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4821598"
/lab_host="DH10B"
/clone_lib="NIH_MGC_97"

FEATURES
source

/note="Organ: testis; Vector: pBluescriptR (modified pBluescript KS+); Site_1: BamHI; Site_2: SalI-XhoI (gtcgag); Oligo-dr primed using primer 5'-TTTTTTTTTTTTTTVN-3', size-selected for average insert size 2.2 kb and normalized to ROT 5. This is a primary library enriched for full-length clones and constructed using the Cap-trapper method (Carninci, in preparation). Library constructed by M. Brownstein (NIMH/NHGRI, National Institutes of Health). Note: this is a NIH_MGC Library."

ORIGIN

Alignment Scores:
 Pred. No.: 1.18e-106 Length: 777
 Score: 1233.00 Matches: 243
 Percent Similarity: 98.39% Conservative: 1
 Best Local Similarity: 97.98% Mismatches: 3
 Query Match: 50.18% Indels: 2
 Gaps: 0
 DB:

US-09-886-954A-1 (1-477) x BG717034 (1-777)

QY 1 MetThrProGluAspProGluThrGlnProLeuLeuGlyProGlyGlySerAla 20
 Db 36 ATGACGCCCGGAGACCCAGAGAAACCCAGCCGCTTCGGGGCTCTCTGGCGGCGCGCG 95
 QY 21 ProArgGlyArgArgValPheLeuAlaAlaPheAlaAlaLeuGlyProLeuSerPhe 40
 Db 96 CCCCCGGCGCGCGCTCTCTCGCGCGCTTCGGCGCTTCGGCGCTTCGGCGCTTCGACGCTTC 155
 QY 41 GlyPheAlaLeuGlyTyrSerSerProAlaIleProSerLeuGlnArgAlaAlaProPro 60
 Db 156 GGGTTCGGCGCTCGCTACAGCTCCCGCGCCATCCCTAGCTCGAGCGCGCGCGCGCGCG 215
 QY 61 AlaProArgLeuAspAlaAlaAlaSerTrpPheGlyAlaValValThrLeuGlyAla 80
 Db 216 GCCCGCGCGCTGACGACGCGCGCGCTCTCTGGTTCGGGGCTTCGTGACCTCGGTGCG 275
 QY 81 AlaAlaGlyValLeuGlyGlyTrpLeuValAspArgAlaGlyArgGlySerLeu 100
 Db 276 GCGCGCGGGGAGTGTCTGGCGCGCTGGCTGGTGGACCGCGCGCGGGCGAAGCTGAGCCTC 335
 QY 101 LeuLeuCysSerValProPheValAlaGlyPheAlaValIleThrAlaAlaGlnAspVal 120
 Db 336 TTGCTGTCTCCGTCGCTTCGTGGCGCGCTTCGTGCTCATCACCGCGCGCGCGCGCGCG 395
 QY 121 TrpMetLeuGlyGlyArgLeuLeuThrGlyLeuAlaCysGlyValAlaSerLeuVal 140
 Db 396 TGGATGCTGTGGGGGCGCGCTCTCTCACCGCGCTGGCGCTGGCGGTTCCTCCCTAGTG 455
 QY 141 AlaProValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuGlySerCys 160
 Db 456 GCCCGCGGTACATCTCCGAAATCGGCTACCCAGCAGTCCGGGGGTTCGTGCTCTCTGT 515
 QY 161 ValGlnLeuMetValValGlyIleLeuLeuAlaTyrIleLeuAlaGlyTrpValLeuGlu 180
 Db 516 GTGCAGCTATGTGTGTGTGGCATCTCTCTGGCTTACTGGCGAGCTGGGTGTGGAG 575
 QY 181 TrpArgTrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCys 200
 Db 576 TGGCGCTGGGTGTGTGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGGGTGG 635
 QY 201 PheMetProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAla 220
 Db 636 TTCATGCCCCGAGACCCCGCGCTTCCTGCTGACTCAGCACAGCGCGCGAGGCGCATGGCC 695
 QY 221 AlaLeuArgPheLeu-TrpGlySerGluGlnGlyTrpGluAspProProIleGlyAlaG 240
 Db 696 GCCCTGCGGTTCCTGTGGGGCTCCGAGCATGCTGGGAGAGACCCCGCATC-GGAGCTGA 754
 QY 240 uGlnSerPheHisLeuAlaLeu 247
 Db 755 CCAGAGCTTCACCTGGACCTG 776

RESULT 15

BI757409
 LOCUS 889 bp mRNA linear EST 25-SEP-2001
 DEFINITION 603029344F1 NIH_MGC_114 Homo sapiens cDNA clone IMAGE:519778 5', mRNA sequence.
 ACCESSION BI757409
 VERSION BI757409.1 GI:15748987
 KEYWORDS EST.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 889)
 NIH-MGC http://mgi.nci.nih.gov/
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Life Technologies, Inc.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone Distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM11500 row: 9 column: 11
 High quality sequence stop: 877.
 Location/Qualifiers
 1. 889
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="taxon:9606"
 /clone="IMAGE:519778"
 /lab_host="DH10B"
 /clone_lib="NIH_MGC_114"
 /note="Organ: Brain; Vector: pCMV-SPORT6; Site_1: NotI; Site_2: EcoRV (destroyed); RNA source anonymous pool of 6 male brains, age range 23-27 yo. Library is oligo-dr primed and directionally cloned (EcoRV site is destroyed upon cloning). Average insert size 1.5 kb, insert size range 1-3 kb. Library is normalized and enriched for full-length clones and was constructed by C. Gruber (Invitrogen). Research Genetics tracking code 019. Note: this is a NIH_MGC Library."

FEATURES

source

ORIGIN

Alignment Scores:
 Pred. No.: 2.48e-106 Length: 889
 Score: 1230.50 Matches: 265
 Percent Similarity: 78.87% Conservative: 0
 Best Local Similarity: 78.87% Mismatches: 4
 Query Match: 50.08% Indels: 69
 Gaps: 1
 DB:

US-09-886-954A-1 (1-477) x BI757409 (1-889)

QY 143 ValTyrIleSerGluIleAlaTyrProAlaValArgGlyLeuLeuGlySerCysValGln 162
 Db 38 GTCTACATCTCGAAATCGCTTACCAGACATCCCGGGGTTCCTCGGCTCTGTGTGCAG 97
 QY 163 LeuMetValValGlyIleLeuLeuAlaTyrIleLeuAlaGlyTrpValLeuGluTrpArg 182
 Db 98 CTAATGATCGTCTCGGCATCTCTCTGGCTTACCTGGCA----- 136
 QY 183 TrpLeuAlaValLeuGlyCysValProProSerLeuMetLeuLeuMetCysPheMet 202
 Db 136 ----- 136
 QY 203 ProGluThrProArgPheLeuLeuThrGlnHisArgArgGlnGluAlaMetAlaLeu 222
 Db 136 ----- 136

```
QY 223 ArgPheLeuTrpGlySerGluGlnGlyTrpGluAspProPheGlyAlaGluGlnSer 242
Db 137 -----GAG 139
QY 243 -PheHisLeuAlaLeuLeuArgGlnProGlyIleTyrlYsProPheIleIleGlyValSe 262
Db 140 CTTTCACTGGCCCTCTCGGGAGCCCGGCATCTACAAGCCCTTCATCATCGCGCTCTC 199
QY 262 rLeuMetAlaPheGlnGlnLeuSerGlyValAsnAlaValMetPheTyAlaGluThrIl 282
Db 200 CCTGATGGCTTCCACGAGCTGTGGGGGTCAACGCCGTCACTGTTCTATGCAGAGACCAT 259
QY 282 ePheGluGluAlaLysPheIysAspSerSerLeuAlaSerValValIleGlyValIleGl 302
Db 260 CTTTGAAGAGGGCAAGTTCAAGGACAGAGCCCTGGCTCGCTCGTGGTGTCTATCCA 319
QY 302 nValLeuPheThrAlaValAlaLeuIleMetAspArgAlaGlyArgArgLeuLeuLe 322
Db 320 GGTGCTGTTCAACAGCTGTGGGGCTCTCATATGAGCAGAGAGCGGAGGCTGCTCCT 379
QY 322 uValLeuSerGlyValValMetValPheSerThrSerAlaPheGlyValAlaTyrlPheLysLe 342
Db 380 GGTCTGTGTCAGGTGTGTCATGTGTTTCAGCAGAGTGCCTTCGGCGCCTACTTCAAGCT 439
QY 342 uThrGlnGlyGlyProGlyAsnSerSerHisValAlaIleSerAlaProValSerAlaGl 362
Db 440 GACCCA-GGTGGCCCTGGCAACTCTCTGCACGTGGGCATCTCGGCGCTGTCTCTGCACA 498
QY 362 nProValAspAlaSerValGlyLeuAlaTrpLeuAlaValGlySerMetCysLeuPheIl 382
Db 499 GCTGTGTAGTCCAGCGTGGGGCTGGCCCTGGCTGGCCGTGGGCGAGCATGTGCTCTTCAT 558
QY 382 eAlaGlyPheAlaValGlyTrpGlyProIleProTrpLeuLeuMetSerGluIlePhePr 402
Db 559 CGCGCGCTTTCGGTGGGTGGGGGCCCATCCCTGGCTCCTCATGTCTAGAGATCTTCCC 618
QY 402 oLeuHisValLysGlyValAlaAlaThrGlyIleCysValLeuThrAsnTrpLeuMetAlaPh 422
Db 619 TCTGCATGTCAAGGGCGTGGCGACAGGCATCTCGCTCCTCAGCAACTGGGCTCATGGCCTT 678
QY 422 eLeuValThrLysGluPheSerSerLeuMetGluValLeuArgProTyrlGlyAlaPheTr 442
Db 679 TCTGTGACCAAGGAGTTCAGCAGCCTCATGGAGGTCCTCAGGCCCTATGGAGCCTTCTG 738
QY 442 pLeuAlaSerAlaPheCysIlePheSerValLeuPheThrLeuPheCysValProGluTh 462
Db 739 GCTTGCTCCG--TTTCTGATCTTCAGTGTCTCTTTCACTTTGTCTGTCTGCTGAAAC 796
QY 462 rLysGlyLysThrLeuGluGlnIleThrAlaHisPheGluGlyArg 477
Db 797 TAAAGGAAGAGACTCTGGAACAATCACAGCCCATTTTGGAGGGCGGA 842
```

Search completed: February 25, 2005, 21:31:27
Job time : 3621 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: February 24, 2005, 05:53:07 ; Search time 43 Seconds
(without alignments)
828.084 Million cell updates/sec

Title: US-09-886-954A-1
Perfect score: 2457
Sequence: 1 MTPDEPTEQPLGPPGSA.....CVPETKGTLEQITAHFEGR 477

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: *
1: /cgn2_6/prodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/prodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/prodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/prodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/prodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/prodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1215.5	49.5	246	4	US-09-949-016-9610
2	615	25.0	488	4	US-10-162-012-46
3	605	24.6	514	4	US-09-489-039A-11902
4	578.5	23.5	476	4	US-09-489-039A-11933
5	575.5	23.4	494	2	US-09-031-392-5
6	575.5	23.4	494	3	US-09-299-549-5
7	575.5	23.4	494	3	US-09-610-417-5
8	570.5	23.2	496	4	US-10-146-704-3
9	568	23.1	493	2	US-09-031-392-10
10	568	23.1	493	3	US-09-299-549-10
11	548.5	22.3	518	4	US-09-610-417-10
12	548.5	22.3	518	4	US-09-679-686B-23
13	542.5	22.1	517	4	US-09-679-686B-18
14	536	21.8	509	2	US-09-031-392-6
15	536	21.8	509	3	US-09-299-549-6
16	536	21.8	509	3	US-09-610-417-6
17	535	21.8	514	4	US-09-679-686B-22
18	529.5	21.6	502	4	US-09-679-686B-2
19	526.5	21.4	519	4	US-09-679-686B-24
20	516.5	21.0	518	4	US-09-679-686B-16
21	516	21.0	584	3	US-08-928-692-13
22	516	21.0	584	2	US-09-339-972-13
23	513	20.9	524	2	US-08-928-692-12
24	513	20.9	524	3	US-09-339-972-12
25	507.5	20.7	501	4	US-09-489-039A-11731
26	506.5	20.6	513	3	US-09-291-922-20
27	498.5	20.3	492	2	US-08-355-844-3

28	498.5	20.3	492	5	PCT-US95-16126-3	Sequence 3, Appli
29	497	20.2	549	3	US-09-291-922-30	Sequence 30, Appl
30	495	20.1	863	4	US-09-894-927B-9	Sequence 9, Appli
31	493	20.1	558	4	US-09-949-016-10630	Sequence 10630, A
32	489	19.9	521	4	US-09-489-039A-9549	Sequence 9549, Ap
33	488.5	19.9	534	2	US-09-031-392-4	Sequence 4, Appli
34	488.5	19.9	534	3	US-09-299-549-4	Sequence 4, Appli
35	488.5	19.9	534	3	US-09-610-417-4	Sequence 19, Appl
36	486.5	19.8	510	4	US-09-679-686B-19	Sequence 10, Appl
37	485.5	19.8	488	2	US-08-928-692-10	Sequence 10, Appl
38	485.5	19.8	488	3	US-09-339-972-10	Sequence 10, Appl
39	478	19.5	294	4	US-09-270-767-45902	Sequence 45902, A
40	478	19.5	562	4	US-10-162-012-44	Sequence 44, Appl
41	474	19.3	329	4	US-09-710-279-1942	Sequence 1942, Ap
42	467.5	19.0	506	4	US-09-248-796A-20075	Sequence 20075, A
43	465.5	18.9	539	3	US-09-291-922-26	Sequence 26, Appl
44	464.5	18.9	523	3	US-09-291-922-24	Sequence 24, Appl
45	458.5	18.7	529	3	US-09-291-922-28	Sequence 28, Appl

ALIGNMENTS

RESULT 1
US-09-949-016-9610
; Sequence 9610, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; FILE REFERENCE: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; CURRENT FILING DATE: 2000-04-14
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 9610
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Human
US-09-949-016-9610

Query Match 49.5%; Score 1215.5; DB 4; Length 246;
Best Local Similarity 99.6%; Pred. No. 8.8e-107;
Matches 241; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
QY 86 LGGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDVWMLGGRLITGLACGVASLAPVYI 145
DB 1 LGGWLVDRAGRKLSLLCSVPFVAGFAVITAAQDVWMLGGRLITGLACGVASLAPVYI 60
QY 146 SSIAYPAVAGLLGSCVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPPLMLLLMCFMPET 205
DB 61 SSIAYPAVAGLLGSCVQLMVMVVGILLAYLAGWLEWRWLAIVLGCVPPLMLLLMCFMPET 120
QY 206 PRFLTQHRQRQBAAMALRFLWGSEQWEDPPIGAE-QSFHALLRPGYKPIIGVSLM 264
DB 121 PRFLTQHRQRQBAAMALRFLWGSEQWEDPPIGAEQSFHALLRPGYKPIIGVSLM 180
QY 265 AFQQLSGVNAVNFYAEITFEFAKFDSSLASVVGVIQVLTAVAAALMDRAGRLLVL 324
DB 181 AFQQLSGVNAVNFYAEITFEFAKFDSSLASVVGVIQVLTAVAAALMDRAGRLLVL 240
QY 325 SG 326
DB 241 SG 242

RESULT 2
US-10-162-012-46
; Sequence 46, Application US/10162012
; Patent No. 6682597
; GENERAL INFORMATION:
; APPLICANT: Curtis, Rory A.J.
; APPLICANT: Silos-Santiago, Inmaculada
; APPLICANT: Gu Wei
; TITLE OF INVENTION: NOVEL HUMAN ION CHANNEL AND TRANSPORTER FAMILY MEMBERS
; FILE REFERENCE: 10448-190001
; CURRENT APPLICATION NUMBER: US/10/162,012
; CURRENT FILING DATE: 2002-06-04
; PRIOR APPLICATION NUMBER: US 60/209,845
; PRIOR FILING DATE: 2000-06-06
; PRIOR APPLICATION NUMBER: US 09/875,321
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: PCT/US01/18340
; PRIOR FILING DATE: 2001-06-06
; PRIOR APPLICATION NUMBER: US 60/209,257
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/875,423
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18398
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/209,238
; PRIOR FILING DATE: 2000-06-05
; PRIOR APPLICATION NUMBER: US 09/928,530
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: PCT/US01/18247
; PRIOR FILING DATE: 2001-06-05
; PRIOR APPLICATION NUMBER: US 60/226,770
; PRIOR FILING DATE: 2000-08-21
; PRIOR APPLICATION NUMBER: US 09/934,421
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: PCT/US01/26096
; PRIOR FILING DATE: 2001-08-21
; PRIOR APPLICATION NUMBER: US 60/279,281
; PRIOR FILING DATE: 2001-03-28
; PRIOR APPLICATION NUMBER: US 10/109,029
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: PCT/US02/09728
; PRIOR FILING DATE: 2002-03-28
; PRIOR APPLICATION NUMBER: US 60/290,288
; PRIOR FILING DATE: 2001-05-11
; PRIOR APPLICATION NUMBER: US (not assigned)
; PRIOR FILING DATE: 2002-05-13
; NUMBER OF SEQ ID NOS: 48
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 46
; LENGTH: 488
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: consensus sequence
US-10-162-012-46
Query Match 25.0%; Score 615; DB 4; Length 488;
Best Local Similarity 35.0%; Pred. No. 1.4e-49;
Matches 179; Conservative 74; Mismatches 168; Indels 90; Gaps 18;
QY 30 AFAAALGPLSGFALGYSSPAIP-----SLQRAAPPAPRLDDAAASFGAVVT 77
Db 2 ALVAALGG---GFLFGYDTGVGGFLALIDFLRFGLLTSSGALAEVLGVSTVLTGLWS 58
QY 78 ---LGAAGCVLGVMLVDRAGKLSLLCSVPFVAGFAVITAAG-----DWW---MLGG 126
Db 59 IFPLGLRGLSLFAGKLGDRFGRKRSLLIALVLFVIGALLSGAAGPYTTIGLWAFYLLIVG 118

QY 127 RLITGLACGVASLVAPVYISEIAYPAVRGLGSCVOLMVVVGILLAYLAGWVL----- 179
Db 119 RVLVLGVGGASVLPVPMYISEIAPKALRGALGSLYQLAITIGLVAAILGLGNKTNDS 178
QY 180 ---EWRWLAVLG--CVPPSLMLLLCMFMPETPRFLLTQHRROEAAMALRFLWNGSEQ--GWE 233
Db 179 ALNSWGRIPILGLQVLPALLLIGLLFLPESPRVLVEKLEAREAVLAKLRGVEDVDQE 238
QY 234 DDPGAEQSFHIA-----LLR---QPGIYKPFIIIGVSLMAFOOLSGVNAVMFYA 279
Db 239 IOEIKAELEATVSEEEKAGKASWGEFRGTRPKVRQRLMGMVNLQAFQOLTGINAIFYS 298
QY 280 ETIFEEAKPKD---SSLASVVVGVIOVLFATAAL--IMDRAGRRLLLVLGGVVMVFTSA 335
Db 299 PTIFKSVGVSDSVASLLVTIIIGVNVFVFTFALLFLVDRCRRPRLLLGAAGMAICFLI 358
QY 336 FGAYFKLTGGPGNSSHVAISAPVSAQPDVDSVGLAWL--AVGSMCLFIAGFVAGHGPWP 394
Db 359 LGA-----SIGVALLLNKPKDPSKAAIGVAVIFILLFTIAFFALGWPWP 405
QY 395 LMSSEIFPLHVKGAVTGICVLTNMLMAFLVTKFSSLMELVRPY--GAFWLA---SAFCI 449
Db 406 VILSELFPPTKVRSKALALATAANWLANFII-----GFLPYITGALGLAGGVVFLV 457
QY 450 FS-----VLTLFCVCPETKGTLEQITAHF 474
Db 458 FAGLLVLFILFVFFVFPETKGTLEBIEELF 488
RESULT 3
US-09-489-039A-11902
; Sequence 11902, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 11902
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-11902
Query Match 24.6%; Score 605; DB 4; Length 514;
Best Local Similarity 33.3%; Pred. No. 1.3e-48;
Matches 156; Conservative 86; Mismatches 187; Indels 40; Gaps 9;
QY 20 APRGER-----VFLAFAAALGPLSGFALGYSSPAIPSLQRAAPPAPRLDDAAASWFGA 74
Db 52 SPRTQRTDRRMNWFVSIAAVAGLLFGLDIGVISGALPFIIDHTFTLSQLQE-----WVVS 107
QY 75 VVTLGAAAGGVVLGWLVDRCRKLKLSLLCSVPFVAGFAVITAAGVWMLLGGRLLTGLAC 134
Db 108 SMMLGAALGALFNGWLSFRLGRKYSLMAGAVLFAVAGSISGSAFAASVEVLLVARVVLGVAV 167
QY 135 GVASLVAPVYISEIAYPAVRGLGSCVOLMVVVGILLAYLAGWVLEW--RWLAVLG--CV 190
Db 168 GIASYTAPLYSEMASENVRGKMSIYQMLVTGLVLAFLSDTAFSYSGNWRMLGVALL 227
QY 191 PPSLMLLLCMFMPETPRFLLTQHRROEAAMALRFLWNGSEQWEDDPIGAESQFHL----- 245
Db 228 PAVILILVFLPNSPRMLAEGRHIEAEVLEMLRDTSEKARDELNIRESLKIKOGGW 287
QY 246 ALLR--OPGIYKPFIIIGVSLMAFOOLSGVNAVMFYAETIFEEAKF---KDSSLASVVVGV 301
Db 288 ALFKINRVRRAVFLGMLLQANQQFTGMNIIIMYAPRIFKMGAGFTTTEQOMATLVVGLT 347

Db 70 FSVGGMIGSVGLFVNRFGRRNSMLVNLAIAGGCLMGFCCKIAESVEMLLILRLIIGL 129
QY 133 ACQVASLVAAPVYISETIAYPAVRGLGSCVQLMWWVVGILLAYLAGWVL---EWRWLAVLG 188
Db 130 FCGLCTGFVPMYIGETSPALRAFGATNLQIGVIGILVAQIFGLKVILGTEDLWPLLLG 189
QY 189 --CVPSLMLLLMCFMPETPRLLTQHRROE-AMAALRFLWSEQGWEDPPICAEOSFH- 244
Db 190 FTILPAILOAALPFCFESPRFLINRKBEBEKAELQRLWGTEDVAQDIQEMKDESMRM 249
QY 245 -----LALLRQPGIYKPFILGVSMLAFQQLSGVNAVNFYAEITFEERAKFKDSSLASV 296
Db 250 SQBKQVTVLELFRVSSYRQPIIISIVLQSLQSGINAVFYSTGIFKDGAVQEPYATI 309
QY 297 VVGVIQVLTAVAAIMDRAGRLLLVLSGVVMVFTSAFGAYFKLTQGGPGNSHVAIS 356
Db 310 GAGVNTVTFVSVFLVERAGRRLHMI-----GLGMAFCSTL 348
QY 357 APVSAQPDVASGLAMLAAGVSMCLFIAGFAVGMPPIWMLMSIIFPLHVKGVATGICVLT 416
Db 349 MTISLLKDNYSNWSFICGAILVFAFPEIGPGPIWFIWELFSGQPRPAAVAGCS 408
QY 417 NWLMAPLVTKFSSLMVLRPYGAFWL-ASAFICFSV-----LFTLFCVPETKGTLEQ 469
Db 409 NMTSNFLV-----GLLPFSATFYLGAYVFTVFLVIFWVFTFFKVPETRGRTFEE 460
QY 470 ITAHFEG 477
Db 461 ITRAFEG 468

RESULT 8

US-10-146-704-3
; Sequence 3, Application US/10146704
; Patent No. 6828421
; GENERAL INFORMATION:
; APPLICANT: Myriad Genetics, Incorporated
; APPLICANT: Sugiyama, Janice
; TITLE OF INVENTION: VAMP-ASSOCIATED PROTEIN A-INTERACTING PROTEINS AND USE THEREOF
; FILE REFERENCE: 1669.01
; CURRENT APPLICATION NUMBER: US/10/146,704
; CURRENT FILING DATE: 2004-05-14
; PRIOR APPLICATION NUMBER: US 60/291,730
; PRIOR FILING DATE: 2001-05-17
; NUMBER OF SEQ ID NOS: 3
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 3
; LENGTH: 496
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-146-704-3

Query Match 23.2%; Score 570.5; DB 4; Length 496;
Best Local Similarity 31.6%; Pred. No. 2.3e-45;
Matches 152; Conservative 88; Mismatches 190; Indels 51; Gaps 9;

QY 26 VFLAFAAALGFLSGFALGYSSPAIPSLQR-----APPAPRLDDAAASWFGAV 75
Db 10 LIFAITVATIGSFQGYNTGVINAPSKIKEFKINTLTDKGNAPPSEVLLTSLWSLSVAI 69
QY 76 VTLGAAGGVLGGLVDRAGRKLSLLICSPVPFVAG---FAVITAAQDVWMLLGGRLITGL 132
Db 70 FSVGGMIGSVGLFVNRFGRRNSMLVNLAVTGGCFLCKVAKSVEMLLILRLIIGL 129
QY 133 ACQVASLVAAPVYISETIAYPAVRGLGSCVQLMWWVVGILLAYLAG--WVL--EWRWLAVLG 188
Db 130 FCGLCTGFVPMYIGETSPALRAFGATNLQIGVIGILVAQIFGLSEILWPLLLG 189
QY 189 --CVPSLMLLLMCFMPETPRLLTQHRROE-AMAALRFLWSEQGWEDPPICAEOSFH- 244
Db 190 FTILPAILOAALPFCFESPRFLINRKBEBEKAELQRLWGTQDVSDIQEMKDESMRM 249

QY 245 -----LALLRQPGIYKPFILGVSMLAFQQLSGVNAVNFYAEITFEERAKFKDSSLASV 296
Db 250 SQBKQVTVLELFRVSSYRQPIIISIVLQSLQSGINAVFYSTGIFKDGAVQEPYATI 309
QY 297 VVGVIQVLTAVAAIMDRAGRLLLVLSGVVMVFTSAFGAYFKLTQGGPGNSHVAIS 356
Db 310 GAGVNTVTFVSVFLVERAGRRLHMI-----GLGMAFCSTL 348
QY 357 APVSAQPDVASGLAMLAAGVSMCLFIAGFAVGMPPIWMLMSIIFPLHVKGVATGICVLT 416
Db 349 MTISLLKDNYSNWSFICGAILVFAFPEIGPGPIWFIWELFSGQPRPAAVAGCS 408
QY 417 NWLMAPLVTKFSSLMVLRPYGAFWL-ASAFICFSV-----LFTLFCVPETKGTLEQ 469
Db 409 NMTSNFLV-----GLLPFSATFYLGAY-VFIIFTGFLITFLAFTFFKVPETRGRTFEDI 467
QY 477 R 477
Db 468 Q 468

RESULT 9

US-09-031-392-10
; Sequence 10, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906
; TELEX: 200154
; INFORMATION FOR SEQ ID NO: 10:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 493 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-392-10

Query Match 23.1%; Score 568; DB 2; Length 493;
Best Local Similarity 32.2%; Pred. No. 3.9e-45;
Matches 159; Conservative 87; Mismatches 160; Indels 88; Gaps 14;
QY 28 LAFAAALGFLSGFALGYSSPAIPSLQRAAPPAPRLDDAAASWFG-----AVVTLGA- 80
Db 16 LAVLIAALGSLFGYGNLVGIN-----APQKVEAFYETWLGKXGEXSPVPTLTL 65
QY 81 -----AAGVILGWLV-----DRAGRKLSLLICSPVPFVAG---FAVITAAQDVWMLL 125
Db 66 WLSVSIFAVGGMIGSVFLVXIGNRLGRKXKXAMLVNVLAIAGLLMGLAKXAXSFEMLIL 125

126 GRLLTGLAGVSLVAPVYISIAIYPAVRGLGSCVQLMVVVGILLAYLAGWVLE-----180
 126 GRFIIGLYCGLSSGVVPMYVGEISPTALRGALGTUNQLGIVIGILIAQVLG--LDSLLGN 183
 181 ---WRWLAIVLGVCPVPSLMLLMCMFETPRFLITQHRREQEAMA--ALRFLWG-----227
 184 ESLWPLLLGLTGVALLQQLLPFCPEPRYLLI--NKNEEARAKKALQRLRGADVSQEV 242
 228 SEQGWEDPPIGAQSFIH--LALLRQPGIYKPFIIIGVSLMAFQOLSGVNAVVFYAEITFEEA 286
 243 AEMKDESRXKXSEKXSVLEFRSRYRQPVIIAIVLQSLQSGINAVFYYSIFEKA 302
 287 KFKDSSLASVVGVIQVLTAVAAALMDRAGRLLLVLSGVVWVSTSAFGAYFKLTQGG 346
 303 GVGQPVYATIGAGVNTVFTVSVFVVERAGRRTLHLL-----G 341
 347 PGNSSHVAISAPVSAQPVDSVGLAWLAVGSMCLFIAGFVAGWGPPIPLLMSEIPLPHVK 406
 342 LGMACCAVMTIALALLDQVPMYSVSIIVAFGVAFVFGPGPIPFIVAELEFSQGP 401
 407 GVATGICVLTNMLMAFLVTKFSSLMELVRPYGAFWLASAFICFISVL-----FTLFCVP 460
 402 PRAIIVAGFSNMTSNTFVGLLFQYIAELGPI-----VFIVFAVLLLLFFIFTFLKVP 454
 461 ETGKLTLEQITAHF 474
 455 ETGKRTFDEIAAAF 468

RESULT 10
 US-09-299-549-10
 ; Sequence 10, Application US/09299549
 ; Patent No. 6136547
 ; GENERAL INFORMATION:
 ; APPLICANT: Tartaglia, Louis A.
 ; APPLICANT: Meng, Xun
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
 ; TITLE OF INVENTION: GLUTEX AND USES THEREOF
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/299,549
 ; FILING DATE: 26-APR-1999
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/031,392
 ; FILING DATE: 26-FEB-1998
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meiklejohn, Ph.D., Anita L.
 ; REGISTRATION NUMBER: 35,283
 ; REFERENCE/DOCKET NUMBER: 07334/072002
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: 617/542-5070
 ; TELEFAX: 617/542-8906
 ; TELEX: 200154
 ; INFORMATION FOR SEQ ID NO: 10:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 493 amino acids
 ; TYPE: amino acid
 ; TOPOLOGY: linear
 ; MOLECULE TYPE: protein
 ; US-09-299-549-10

Query Match 23.1%; Score 568; DB 3; Length 493;
 Best Local Similarity 32.2%; Pred. No. 3.9e-45;
 Matches 159; Conservative 87; Mismatches 160; Indels 88; Gaps 14;
 QY 28 LAFAAALGSLFSGFALGYSSPAIPSLQRAAPAPRLDDAAASWFG-----AVVILGA- 80
 DB 16 LAVLILAAALGSGFYGYNLGVN-----APQKVIKFAFYETWLGXGXPVPTLTL 65
 QY 81 -----AAGVLLGGWLV-----DRAGKLSLLCSPVFPVAG---FAVITAAQDVMWLLG 125
 DB 66 WLSLSVSIFAVGGMISGLVXIGNRLGRKXKAMLVNVLATAGLGLMLAKXKXSFEMLL 125
 QY 126 GRLLTGLAGVSLVAPVYISIAIYPAVRGLGSCVQLMVVVGILLAYLAGWVLE-----180
 DB 126 GRFIIGLYCGLSSGVVPMYVGEISPTALRGALGTUNQLGIVIGILIAQVLG--LDSLLGN 183
 QY 181 ---WRWLAIVLGVCPVPSLMLLMCMFETPRFLITQHRREQEAMA--ALRFLWG-----227
 DB 184 ESLWPLLLGLTGVALLQQLLPFCPEPRYLLI--NKNEEARAKKALQRLRGADVSQEV 242
 QY 228 SEQGWEDPPIGAQSFIH--LALLRQPGIYKPFIIIGVSLMAFQOLSGVNAVVFYAEITFEEA 286
 DB 243 AEMKDESRXKXSEKXSVLEFRSRYRQPVIIAIVLQSLQSGINAVFYYSIFEKA 302
 QY 287 KFKDSSLASVVGVIQVLTAVAAALMDRAGRLLLVLSGVVWVSTSAFGAYFKLTQGG 346
 DB 303 GVGQPVYATIGAGVNTVFTVSVFVVERAGRRTLHLL-----G 341
 QY 347 PGNSSHVAISAPVSAQPVDSVGLAWLAVGSMCLFIAGFVAGWGPPIPLLMSEIPLPHVK 406
 DB 342 LGMACCAVMTIALALLDQVPMYSVSIIVAFGVAFVFGPGPIPFIVAELEFSQGP 401
 QY 407 GVATGICVLTNMLMAFLVTKFSSLMELVRPYGAFWLASAFICFISVL-----FTLFCVP 460
 DB 402 PRAIIVAGFSNMTSNTFVGLLFQYIAELGPI-----VFIVFAVLLLLFFIFTFLKVP 454
 QY 461 ETGKLTLEQITAHF 474
 DB 455 ETGKRTFDEIAAAF 468
 RESULT 11
 US-09-610-417-10
 ; Sequence 10, Application US/09610417
 ; Patent No. 6346374
 ; GENERAL INFORMATION:
 ; APPLICANT: Tartaglia, Louis A.
 ; APPLICANT: Meng, Xun
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
 ; TITLE OF INVENTION: GLUTEX AND USES THEREOF
 ; NUMBER OF SEQUENCES: 10
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Fish & Richardson P.C.
 ; STREET: 225 Franklin Street
 ; CITY: Boston
 ; STATE: MA
 ; COUNTRY: USA
 ; ZIP: 02110-2804
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Diskette
 ; COMPUTER: IBM Compatible
 ; OPERATING SYSTEM: Windows95
 ; SOFTWARE: FastSeq for Windows Version 2.0
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/610,417
 ; FILING DATE: 05-Jul-2000
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 09/299,549
 ; FILING DATE: <Unknown>
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Meiklejohn, Ph.D., Anita L.
 ; REGISTRATION NUMBER: 35,283

Query Match 22.1%; Score 542.5; DB 4; Length 517;
Best Local Similarity 29.9%; Pred. No. 1.1e-42;
Matches 150; Conservative 88; Mismatches 202; Indels 61; Gaps 14;

QY 17 GCSAPRR---VFLAFAAALGPLSPFALGYSPAL---PSLQRAAPPA-----61
DB 11 GCKDPGRLLTLVFFTCVVAATGGLIFGDIIGISGVTSNMNPLFKKFFPEVDYKQMKGS 70
QY 62 ---PRLDDAASWFGAVTTLGAAGGVLGWLVDRAGRKLSLLCSVPFVAGAVITAA 117
DB 71 ASQCKYDNLQLOTTFTSSLYLAALVSSFFAAITRVVGRKSMFTGGTLFLGAALNGAA 130
QY 118 QDVMLLGLRLTLGLACGVASLAPVYISEIAPVAVRGLLSCVQLMVVVGILLAYL---174
DB 131 ENIAMLIIVGRILLGVGVGFANGSVPYLSEMAPARLGMNLNGFOLMITIGILAAALINY 190
QY 175 -----AGVLEWRMLAVLGCVPPLMLLLCMFMPETPRFLTQHRROEAMALRFLMGS 228
DB 191 DTNKIKAGY--GWRISLAIAAIVPAGIITLGSFFLPTDPTNSLTERGHPAARMLNIRGS 248
QY 229 E---QGWDPPIGAEQSPHALLROP-----GIYKP-FIIGVSLMAFQQLSGVNAVWF 277
DB 249 DVDISEYADLVVASEES---KLQHPWENILQKRPOLTNWAIMPPFQQTGINVIMF 305
QY 278 YAEETIPEEAKFK-DSSLAS-VVVGIVQLVTAALIMDRGRRLLLVLSGVVWVFSTSA 335
DB 306 YAPVLFETLGFKGDASLSMAVITGLVNVFATLVSFTVDRGRRLFLQGGTQMLLSQLV 365
QY 336 FGAYFKLTGGGPGNSHVAISAPVSAQPDASVGLAWLAVGSMCLFIAGFVAGWGPPIW 395
DB 366 VGTLLAVKFGTSG-----VGEMP---KGYAAAVALFICLYVAGFAWSWGPGLWL 411
QY 396 LMSEIFPLHVKGVATGICVLTNWMAFLVTKFSSLMELVRLPYGAFWLASAFICISVLFT 455
DB 412 VSEIFPLEIRPAGOSINVSNNLFTFVIAQAFLTNLCHK-FGLFYFAGVWVIMTVEI 470
QY 456 LFCVPETKTKLEQITAHFEG 476
DB 471 ALFLPETKNVPIEMVWVKG 491

RESULT 14

US-09-031-392-6
; Sequence 6, Application US/09031392
; Patent No. 5942398
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES
; TITLE OF INVENTION: ENCODING GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/031,392
; FILING DATE: 26-FEB-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Meiklejohn, Ph.D., Anita L.
; REGISTRATION NUMBER: 35,283
; REFERENCE/DOCKET NUMBER: 07334/072001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 617/542-5070
; TELEFAX: 617/542-8906

TELEX: 200154
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 509 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-09-031-392-6

Query Match 21.8%; Score 536; DB 2; Length 509;
Best Local Similarity 30.5%; Pred. No. 4.3e-42;
Matches 159; Conservative 86; Mismatches 172; Indels 104; Gaps 16;

QY 13 LGPPGCSAPRR---RVFLAFAAALGPLSPGFA-----LGVSSPA 50
DB 8 IGSEDEPQORVCTTLVAVSFVLSQFQYINIGVINAQKVTEQSYNETWLRQGP 67
QY 51 IPSLQRAAPPRLDDAAASWFGAVVTILGA-----AAGVGLGWL---DRGRKLSL 100
DB 68 GPS---STPP-----GTLTTLWALSVAIFSVMSSFLIGISQWLGRKRAM 112
QY 101 LLCSPVFFVAG---FAVITAAQDVWMLLQGRLLTGLACGVASLAPVYISEIAPVAVRGLL 157
DB 113 LVNNVAVLGGSLMGLANAAASYEMLILGRFLIGAYSGLTSLVPMYVGEIAPTHLRGAL 172
QY 158 GSCVQLMVVVGILLAYLAGWLE-----WRMLAVLGCVPPLMLLLCMFMPETPRFL 209
DB 173 GTLQOLAIVIGILLTAQVLG--LESLLGTASLWPLLLGLTULPALLQLVLLPFCPSRYL 230
QY 210 -LTQHRROEAMALRFL--WGSEQ-----WEDPPFAGSQSFHALLRQPGI---YKPTII 259
DB 231 YIQNLGEPARKSLKRLTGWADVSVGLAELDKERKLERERPLSLIQLLGSRTHQPLII 290
QY 260 GYSLMAFQQLSGVNAVMEFYAETIPEEAKFKDSSLASVVVGVIVLFTAVAALIMDRAGR 319
DB 291 AVVLQLSQQLSGINAVFYISTSIPTAGVGOPAVATIGAGVNTVFTLVSVLLVERAGR 350
QY 320 LLLVLSGVVWVFSTSAFGAYFKLTGGGPGNSHVAISAPVSAQPDASVGLAWLAVGSMC 379
DB 351 TLHLL-----GLAGCGCAILMTVALLLLRVPMSYVSIIVAF 389
QY 380 LFIAGFVAGWGPPIWMLMSEIFPLHVKGVATGICVLTNWMAFLVTKFSSLMELVRLPYG 439
DB 390 GFVAFFEIGPGPIWFIWELFVQGPRAAMAVAGFSNWTNFTIGMGFQVVAEAMGPV- 448
QY 440 AFWLASAFICISVL-----FTLCVDETCKTLEQITAHF 474
DB 449 -----VELLFAVLLGLGFIIFLRVPETRGRTFDQISAAF 483

RESULT 15

US-09-299-549-6
; Sequence 6, Application US/09299549
; Patent No. 6136547
; GENERAL INFORMATION:
; APPLICANT: Tartaglia, Louis A.
; APPLICANT: Weng, Xun
; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING
; TITLE OF INVENTION: GLUTEX AND USES THEREOF
; NUMBER OF SEQUENCES: 10
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson P.C.
; STREET: 225 Franklin Street
; CITY: Boston
; STATE: MA
; COUNTRY: USA
; ZIP: 02110-2804
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/09/299,549
FILING DATE: 26-APR-1999
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/031,392
FILING DATE: 26-FEB-1998
ATTORNEY/AGENT INFORMATION:
NAME: Meiklejohn, Ph.D., Anita L.
REGISTRATION NUMBER: 35,283
REFERENCE/DOCKET NUMBER: 07334/072002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 617/542-5070
TELEFAX: 617/542-8906
TELEX: 200154
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 509 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: protein
US-09-299-549--6

Query Match 21.8% Score 536; DB 3; Length 509;
Best Local Similarity 30.5%; Pred. No. 4.3e-42;
Matches 159; Conservative 86; Mismatches 172; Indels 104; Gaps 16;

QY	13	LGPPGGSAPRGR---RVFLAAFAAALGPLSGFPA-----LGYSSPA 50
DB	8	IGSEDEPPQORVTGLVLAFAVSLGSLQFGYNIGVINAPQKVIQSYNETWLGKQGE 67
QY	51	IPSLQRAAPPAPRLDDAASWFGAVTLGA-----AAGVLGGLV-----DRAGKLSL 100
DB	68	GPS---SIPP-----GTLTTLWALSVAIFSVGGMISSFLIGIISQWLGRKRAM 112
QY	101	LLCSVPFVAG---FAVITRAQDVVMLLGGRLTLGLAGVASLVAPVYVISEIAYPAVRGLL 157
DB	113	LNNVLAVLGGSLMGLANAASYLEMLILGRFLIGAYSLTSGLVPMVYVGEIAPTHRGAL 172
QY	158	GSCVQLMVMVGIILLAVLAGHVLE-----WRLAVLGCVPSPSLMLLLMCFMPTPREFL 209
DB	173	GTNLQLAIVIGILIAQVLG--LESLLGTASLWPLLGLTVLPALLQLVLLPFCSPRYL 230
QY	210	LTQHRROEAMALRFL--WGSEGG-----WEDPPIGAEQSFHLALLRQPGI----YKPFII 259
DB	231	YIIQNLGEPARKSLKRLTGWADVSGVLAELDKERLERPLSLQLLGSRTHRQPLII 290
QY	260	GVSLMAFQQLSGYNAMVFAETIFERAKFKDSSLASVVVGIVQVLTAVAALIMDRAGR 319
DB	291	AVVLQSLQQSGINAVFYSTSFETAGVGPAYATIGAVNTVFTLVSVLLVERAGR 350
QY	320	LLVLSGVMVVFSTSAFGAYFKLTQGGPGNSSHVAISAPVSAQPDVDSVGLAWLAVGSMC 379
DB	351	TLHLL-----GLAGMCGCAILMTVALLLRLRPVPMYSIVAI 389
QY	380	LFTAGFVAGMPDPWLLMSEIPFLHVKGVATGICVLTNNLMALFVTKBSSLMELVRPYG 439
DB	390	GFVAFFEIGPGIPWETIVAELEFSGGPRPAMAVAGFSNMTSNETIIGMGFOYVAEAMGPY- 448
QY	440	AFWLASAFICFVSL-----FTLFCVPETKGTLEQITAHF 474
DB	449	-----VFLFVALLLGGFFITFLRVPETKGRTFDQISAAP 483

Search completed: February 24, 2005, 06:05:26
Job time : 45 secs

THIS PAGE BLANK (USPTO)